

GenCore version 5.1.4.p5.4578
nt (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: April 11, 2003, 14:21:02 ; Search time 17 Seconds
 (without alignments) 486.327 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459

Sequence: 1 LSPAPFPFOYCVETDIPKLT.....SSGDQFLSVSPFTWSKPKP 86

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :
 1: PIR_73: *
 2: PLR1: *
 3: PLR2: *
 4: PLR3: *
 4: PLR4: *

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	459	100.0	1065	2 A57410	transcription fact
2	430	93.7	1075	2 A57377	transcription fact
3	114	24.8	716	2 S45262	Nr-AT component -
4	96.5	21.0	718	2 JC5805	transcription fact
5	81.5	17.8	329	2 G86253	hypothetical prote
6	73.5	16.0	3488	2 T34418	hypothetical prote
7	72	15.7	2130	2 AB0821	probable exported
8	71.5	15.6	792	2 JC7122	protein kinase (EC
9	70	15.3	643	2 S5435	permease-like prot
10	70	15.3	649	1 B37953	transcription regu
11	70	15.3	652	2 G95177	hypothetical prote
12	70	15.3	652	2 E98044	hypothetical prote
13	70	15.3	660	2 G97912	hypothetical prote
14	70	15.3	661	2 E95042	hypothetical prote
15	69.5	15.1	638	2 B35816	transcription regu
16	69	15.0	649	2 A35816	transcription regu
17	67.5	14.7	314	2 I38864	transcription regu
18	67	14.6	177	2 H90631	probable fibrillar
19	67	14.6	177	2 G85482	probable fibrillar
20	67	14.6	1112	2 T32733	transcription regu
21	66.5	14.5	1039	2 T28905	hypothetical prote
22	66	14.4	728	1 A60185	hepatocyte growth
23	66	14.4	1944	2 A59438	hepatocyte growth
24	65.5	14.3	652	2 S47979	transcription regu
25	65	14.2	635	2 T20587	hypothetical prote
26	64.5	14.1	421	2 I49734	HNF-3/foxa-head ho
27	64	13.9	84	2 S76052	hypothetical prote
28	64	13.9	616	2 S38060	carboxylic acid tr
29	64	13.9	728	1 A35644	hepatocyte growth

30	63	13.7	498	2 T09436	gag polyprotein -
31	63	13.7	570	2 S47142	matrig type A prot
32	63	13.7	570	2 T49181	cyclophilin-like p
33	63	13.7	1216	2 AH1535	pyruvate-Flavodoxi
34	62.5	13.6	474	2 T39587	serine/Chreonine
35	62.5	13.6	671	2 S61693	probable membrane
36	62.5	13.6	1530	2 E82085	glutamate synthase
37	62.5	13.5	294	2 G97484	dihydrodipicolinat
38	62	13.5	294	2 AG2702	hypothetical prote
39	62	13.5	579	2 AF2094	hypothetical prote
40	62	13.5	963	2 T40290	AND-1 protein - Af
41	62	13.5	1127	2 T30334	immunoglobulin A1
42	62	13.5	1856	2 C95008	probable serine es
43	61.5	13.4	279	2 C75491	hypothetical prote
44	61.5	13.4	441	2 B86252	probable receptor
45	61.5	13.4	838	2 A96557	

ALIGNMENTS

RESULT 1

A57410

transcription factor NFATC3 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 08-Feb-1996

C:Accession: A57410

R:Ho, S.N.; Thomas, D.J.; Timmerman, L.A.; Li, X.; Francke, U.; Crabtree, G.R.

J. Biol. Chem. 270, 19898-19907, 1995

A:Title: NFATC3, a lymphoid-specific NFATC family member that is calcium-regulated an

A:Reference number: A57410; MUID:95378239; PMID:7650004

A:Accession: A57410

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1065 <HOA>

A:Cross-references: GB:U28807

Query Match 100.0%; Score 459; DB 2; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 5.8e-42;
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSPAPFPFOYCVETDIPKLTSEDOAILPGKLEICSDDOGNLSPRSRTSVDDGLGSO 60
 Db 311 LSPAPFPFOYCVETDIPKLTSEDOAILPGKLEICSDDOGNLSPRSRTSVDDGLGSO 370

QY 61 YPLKDDSSGDQFLSVSPFTWSKPKP 86
 Db 371 YPLKDDSSGDQFLSVSPFTWSKPKP 396

RESULT 2

A57377

transcription factor NFATX - human

C:Species: Homo sapiens (man)

C>Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 05-Nov-1999

C:Accession: A57377

R:Masuda, E.S.; Naito, Y.; Tokumitsu, H.; Campbell, D.; Saito, F.; Hannum, C.; Arat,

Mol. Cell. Biol. 15, 2697-2706, 1995

A:Title: NFATX, a novel member of the nuclear factor of activated T cells family that

A:Reference number: A57377; MUID:95257951; PMID:7739550

A:Accession: A57377

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1075 <MAS>

A:Cross-references: GB:U14510; NID:9780373; PIDN:AAA86308.1; PID:9780374

C:Keywords: transcription factor

Query Match 93.7%; Score 430; DB 2; Length 1075;
 Best Local Similarity 91.9%; Pred. No. 8.7e-39;
 Matches 79; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSPAPFPFOYCVETDIPKLTSEDOAILPGKLEICSDDOGNLSPRSRTSVDDGLGSO 60
 Db 111 LSPAPFPFOYCVETDIPKLTSEDOAILPGKLEICSDDOGNLSPRSRTSVDDGLGSO 60

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 14:20:32 ; Search time 13 Seconds
(without alignments)
274.382 Million cell updates/sec

Title: US-09-550-115A-2

Sequence: 1 LSPAPPPGYCVETDIPKLT.....SSGDFLSPVSPFTWSKKPK 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459	100.0	1075	1	NFC3_MOUSE
2	430	93.7	1075	1	NFC3_HUMAN
3	122	26.6	902	1	NFC4_HUMAN
4	114	24.8	943	1	NFC1_HUMAN
5	102	22.2	822	1	NFC1_MOUSE
6	96.5	21.0	717	1	NFC1_PIG
7	71.5	15.6	513	1	NFC1_MOUSE
8	70	15.3	649	1	DMPL_HUMAN
9	70	15.3	652	1	TFE2_MESAU
10	70	15.3	660	1	ALIA_STRPN
11	69	15.0	649	1	TFE2_MOUSE
12	67.5	14.7	314	1	NFC2_MOUSE
13	67	14.6	558	1	NFC2_HUMAN
14	66.5	14.5	486	1	VNHL_MOUSE
15	66	14.4	728	1	RIK3_MOUSE
16	65.5	14.3	326	1	HGF_MOUSE
17	65	14.2	898	1	NIP2_MOUSE
18	64.5	14.1	421	1	CIZ1_HUMAN
19	64.5	14.1	535	1	FXJ1_MOUSE
20	64.5	14.1	708	1	FXJ1_HUMAN
21	64	13.9	728	1	ICAL_MOUSE
22	64	13.9	728	1	JEN1_MOUSE
23	62	13.5	500	1	HGF_MOUSE
24	62	13.5	500	1	HGF_MOUSE
25	61.5	13.4	367	1	SHK3_MOUSE
26	61.5	13.4	2032	1	THA_MOUSE
27	61	13.3	508	1	GAG_MOUSE
28	61	13.3	508	1	GAG_MOUSE
29	61	13.3	1203	1	GAG_MOUSE
30	60.5	13.2	421	1	MGR5_MOUSE
31	60.5	13.2	503	1	FXJ1_MOUSE
32	60.5	13.2	586	1	HGF_MOUSE
33	60.5	13.2	642	1	NA95_MOUSE

34	60.5	13.2	653	1	EGLN_MOUSE
35	60.5	13.2	2895	1	HYD_MOUSE
36	60	13.1	488	1	HHR_MOUSE
37	60	13.1	498	1	HHR_MOUSE
38	60	13.1	529	1	HHR_MOUSE
39	60	13.1	607	1	HHR_MOUSE
40	60	13.1	611	1	HHR_MOUSE
41	60	13.1	646	1	HHR_MOUSE
42	60	13.1	1069	1	ACAA_MOUSE
43	60	13.1	1093	1	ACAA_MOUSE
44	60	13.1	1856	1	SM5_MOUSE
45	59.5	13.0	467	1	YATU_MOUSE

ALIGNMENTS

RESULT 1
ID NFC3_MOUSE STANDARD: PRT: 1075 AA.
AC P97305.060896;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 3 (T cell)
DE transcription factor NFAT3 (NF-AT3) (NF-AT4) (NFATX).
GN NFATC3 OR NFAT4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scuriongnath; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM X1).
RC TISSUE=Thymus;
RX MEDLINE=95378239; PubMed=7650004;
RA Ho S.N., Thomas D.J., Timmerman L.A., Li X., Francke U.,
RA Crabtree G.R.;
RT "NFATC3, a lymphoid-specific NFATC family member that is
RT calcium-regulated and exhibits distinct DNA binding specificity."
RL J. Biol. Chem. 270:19988-19907(1995).
RP [2]
RP SEQUENCE OF 12-1075 FROM N.A. (ISOFORMS X1; X2 AND DELTA-X).
RC TISSUE=Thymic lymphoma;
RX MEDLINE=97170074; PubMed=9017603;
RA Liu J., Koyano-Nakagawa N., Amasaki Y., Saito-Ohe
RA Inai S.-I., Takano T., Arai N., Yokota T., Arai I.
RT "Calcineurin-dependent nuclear translocation of
RT factor NFATX: molecular cloning and functional characterization
RL Mol. Biol. Cell 8:157-170(1997).
RN [3]
RP REVIEW.
RX MEDLINE=99189746; PubMed=10089876;
RA Crabtree G.R.;
RT "Genetic signals and specific outcomes: signaling through Ca2+,
RT calcineurin, and NF-ATX."
RL Cell 96:611-614(1999).
CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 (BY
CC SIMILARITY).
CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
CC CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC5 OR MEMBERS OF THE
CC ACTIVATING PROTEIN-1 FAMILY, MAP, GATA4 AND C/EBP300 CAN ALSO BIND
CC THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND
CC NUCLEAR AFTER ACTIVATION. THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: X1 (SHOWN HERE), X2 AND DELTA-X;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.

SNC

CC -I- TISSUE SPECIFICITY: EXPRESSED IN THYMUS. HEAVILY EXPRESSED IN
CC MUSCLE, SPLEEN AND KIDNEY. ALSO EXPRESSED IN LYMPH NODE.
CC
CC -I- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC COOPERATIVE INTERACTIONS WITH AIF FACTORS. (BY SIMILARITY).
CC
CC -I- PTM: PHOSPHORYLATED BY NPAIC-KINASE; DEPHOSPHORYLATED BY
CC CALCINEURIN (BY SIMILARITY).
CC
CC -I- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```
CC -----  
DR EMBL; D85612; BAI12833.1; -  
DR EMBL; U28807; AAA93249.1; -  
DR MGD; MG1.103296; Nfatc3.  
DR InterPro; IPR002909; IPT_TIG.  
DR InterPro; IPR000451; NF_Rel_dor_fam.  
DR Pfam; PF00554; RHD; 1  
DR Pfam; PF01833; TIG; 1  
DR SMART; SMO0429; IPT; 1  
DR PROSITE; PS01204; REL_1; FALSE_NEG.  
DR PROSITE; PSS0254; REL_2; 1  
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;  
KW Alternative splicing; Repeat; Poly-PRO.  
FT DOMAIN 24 29  
FT DOMAIN 109 114  
FT DOMAIN 207 308  
FT REPEAT 207 223  
FT REPEAT 236 252  
FT REPEAT 292 308  
FT DOMAIN 273 275  
FT DOMAIN 444 451  
FT DOMAIN 686 688  
FT DOMAIN 1031 1040  
FT VARSPLIC 468 497  
FT VARSPLIC 1035 1075  
  
VNFEDRMSQISVSQAIEVMKDITLPAPASDLMTSHSAA
```

FT	CONFLICT	12	34	LDKPIYFGEDGAPAPPSPCSRPA -> PRVLFVSQAOLBS
FT				RTRGPSDL (IN REF. 2).
FT	CONFLICT	57	61	ITPLC -> NSSLG (IN REF. 2).
FT	CONFLICT	68	68	P -> LQ (IN REF. 2).
FT	CONFLICT	83	83	SH -> GY (IN REF. 2).
FT	CONFLICT	90	93	CEIP -> GDIS (IN REF. 2).
FT	CONFLICT	113	113	F -> I (IN REF. 2).
FT	CONFLICT	121	124	QOBL -> HOGT (IN REF. 2).
FT	CONFLICT	132	132	Q -> H (IN REF. 2).
FT	CONFLICT	132	132	F -> Y (IN REF. 2).
FT	CONFLICT	140	140	G -> R (IN REF. 2).
FT	CONFLICT	641	641	V -> A (IN REF. 2).
FT	CONFLICT	646	646	REMDPEVSPLVPHSAQAOPRSSETG -> KERKQICLAL
FT	CONFLICT	707	734	EHLCICLDESPALRLDR (IN REF. 2).
FT				G -> S (IN REF. 2).
FT	CONFLICT	746	746	H -> Q (IN REF. 2).
FT	CONFLICT	1075	1075	
SQ	SEQUENCE	1075 AA:	115450 MW:	4ED38C9A6F52B8 CRC64:

Query Match	100.0%	Score 459:	DB 1:	Length 1075;
Best Local Similarity	100.0%	Pred. NO. 2.1e-41;		
Matches 86; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 LSPAPPFQCYCETDPLPLKTRKTSDDQAILPGKLEICSDDQGNLSPSRKTSVDDGIGSQ 60

Db 321 LSPAPPFQCYCETDPLPLKTRKTSDDQAILPGKLEICSDDQGNLSPSRKTSVDDGIGSQ 380

QY 61 YPLKDDSGDQFLSVSPPTWASKPP 86

Db 381 YPLKDDSGDQFLSVSPPTWASKPP 406

RESULT 2

ID	NFPC3_HUMAN	STANDARD:	PT:	1075 AA.
AC	012968:	014516:	099840:	099842:
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Nuclear factor of activated T-cells, cytoplasmic 3 (T cell transcription factor NFAT4) (NF-ATc3) (NF-AT4) (NFATX).			
GN	NFATC3 OR NFAT4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606:			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS A; B AND X2/C).			
RC	TISSUE=T-cell, and Skeletal muscle;			
RX	MEDLINE=95269130: PubMed=7749981;			
RA	Hoey T., Sun Y.-L., Williamson K., Xu X.;			
RT	"Isolation of two new members of the NF-AT gene family and functional characterization of the NF-AT proteins.";			
RL	Immunoly 2:461-472(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM X1).			
RC	TISSUE=T-cell;			
RX	MEDLINE=95257951: PubMed=7735550;			
RA	Masuda E.S., Naito Y., Tokumitsu H., Campbell D., Saito F., Hannum C.			
RT	Arai K.-I., Arai N.;			
RL	"NFATx, a novel member of the nuclear factor of activated T cells family that is expressed predominantly in the thymus.";			
RN	Mol. Cell. Biol. 15:2697-2706(1995).			
RP	[3]			
RN	SEQUENCE FROM N.A. (ISOFORMS X2; X3 AND X4).			
RC	TISSUE=T-cell, and Fibroblast;			
RX	MEDLINE=98430664: PubMed=9759864;			
RA	Imamura R., Masuda E.S., Naito Y., Imai S.-I., Fujino T., Takano T.,			
RT	Arai K.-I., Arai N.;			
RL	"Carboxy-terminal 15 amino acids sequence of NFATx1 possibly created by tissue specific splicing is conserved among NFAT family proteins and is essential for transactivation activity in T cells.";			
RN	J. Immunol. 161:3455-3463(1998).			
RP	[4]			
RN	SEQUENCE FROM N.A. (ISOFORM X4).			
RX	MEDLINE=99425270: PubMed=1043829;			
RA	Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,			
RT	Fuhmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,			
RL	Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,			
RT	Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;			
RL	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";			
RN	Genomics 60:295-308(1999).			
RN	[5]			
RP	MUTAGENESIS.			
RX	MEDLINE=98292182: PubMed=9630228;			
RA	Zhu J., Shibaaki F., Price R., Guillemot J.-C., Yano T., Doetsch V.,			
RT	Wagner G., Ferrara P., McKee R.;			
RL	"Intramolecular masking of nuclear import signal on NF-AT4 by casein kinase I and MEK1.";			
RN	Cell 93:851-861(1998).			
RN	[6]			
RP	REVIEW.			
RX	MEDLINE=99189746: PubMed=10089876;			
RA	Crabtree G.R.;			
RT	"Genetic signals and specific outcomes: signaling through Ca2+, calcineurin, and NF-AT.";			
RL	Cell 96:611-614(1999).			
RP	[7]			
RN	FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2.			
CC	SUBUNIT: MEMBER OF THE MULTICOMPONENT NFAT TRANSCRIPTION COMPLEX THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE ACTIVATING PROTEIN-1 FAMILY MAP, GATA4 AND CBF/300 CAN ALSO BIND			

The complex NFATC proteins bind to DNA as monomers.

-1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.

-1- ALTERNATIVE PRODUCTS: 6 ISOFORMS; X1 (SHOWN HERE), X2/C, X3, X4, A, AND B; ARE PRODUCED BY ALTERNATIVE SPLICING.

-1- TISSUE SPECIFICITY: ISOFORM X1 IS PREDOMINANTLY EXPRESSED IN THYMUS AND IS ALSO FOUND IN PERIPHERAL BLOOD LEUKOCYTES AND KIDNEY. ISOFORM X2 IS PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLE AND IS ALSO FOUND IN THYMUS, KIDNEY, TESTIS, SPLEEN, PROSTATE, OVARY, SMALL INTESTINE, HEART, PLACENTA AND PANCREAS. ISOFORM X3 IS EXPRESSED IN THYMUS AND KIDNEY. ISOFORM X4 IS EXPRESSED IN THYMUS AND SKELETAL MUSCLE.

-1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND COOPERATIVE INTERACTIONS WITH APC FACTORS.

-1- PTM: PHOSPHORYLATED BY NFATC-KINASE. DEPHOSPHORYLATED BY CALCINEURIN (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or_send_an_email_to_license@isb-slb.ch)

-- --
CC
DR EMBL; LA1067; AAA979174.1; --
DR EMBL; U14510; AAA865308.1; --
DR EMBL; UB5428; AAAB6595.1; --
DR EMBL; UB54329; AAAB6596.1; --
DR EMBL; UB54330; AAAB6597.1; --
DR EMBL; AC004531; AAC27434.1; --
GENEW; HGNC:7777; NFATC3.
MIM; 602698; --
DR InterPro; IPR0042909; IPT_TIG.
DR InterPro; IPR000451; NE_Rel_dor_fam.
DR Pfam; PF00554; RHD; 1.
DR Pfam; PF01833; TIG; 1.
SMART; SMART0439; IPT_1.
PROSITE; PS01204; REL_1; FALSE_NEG.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
RR Alternative splicing; Repeat; Phosphorylation.
FT DOMAIN 24 29 POLY-PRO.
FT DOMAIN 109 114 CALCIUNEURIN-BINDING.
FT DOMAIN 207 308 3 x Sp REPEATS.
FT REPEAT 207 223 1.
FT REPEAT 226 252 2.
FT REPEAT 292 308 3.
FT DOMAIN 273 275 NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 444 451 DNA-BINDING.
FT DOMAIN 686 688 NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 1032 1041 NUCLEAR EXPORT SIGNAL.
VARSPUBLIC 700 1075

VLMQDEHREETDLSVSFLPVHPAQGTQPSSDSGCSHDYS
LSGORSLICSIPGYASNVTSIHLPOLCRDEVSKCEKHI
PSPIVOHFQTPTTPPVGSSSYOPMGONTNVMYCPLPINAA
SGEFDVYLFOODATTGLGVNLGCOPSIFPHSNSTGG
HLNAHTPSHYTLPHLGSMYGCHSWGCGRSLSPPADITGP
OSSOLODIRTGPHSGSATTTAPSAHSRPASPPLSGCTPSP
QLPMPTOSPSSGTASSPSPATRNMHGSHDTAOASTGGGL
SAPSLLICHSLCDPASFPDDATAVISIKPEPDRENENTIG
LDITDLVDNELIQRDMQSIVSOAGAVSRQAPLPSPESLD
LAGRDGI -> GTRSSDILL (in isoform A).
SLDPHPNOTGRPPSSDGCSHDYSVLSCGRSLICSIPGYAS
MTSSHLLPOLCRDEVSKCEKHMI PSPIHVPOTTPPY
GSTYPMDNTNYNGPCPLLINAASSGFDFSVLTPOQDATLS
CLVNLCGPLSSIIPHNSGSCTGHLLNHASHVTTHILO
SMGYHCNSTGRRLSSPADVITGO PSSOLQPIITVGPBSHG
SATTAASAHSNPRIKLASSPGSPPOLPMPPOSPPSSGTASS

[illegible]

```

CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC1 IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC NFATC1 PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, LUNG, KIDNEY,
CC TESTIS AND OVARY. WEAKLY EXPRESSED IN SPLEEN AND THYMUS. NOT
CC EXPRESSED IN PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED BY NFATC KINASE; DEPHOSPHORYLATED BY
CC CALCINEURIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: LA1066; AAA79175.1; -.
CC TRANSFAC: T02462; -.
CC GeneW: HGNC:7778; NFATC4.
CC DR InterPro: IPR002909; IPT_TIG.
CC DR InterPro: IPR000451; NF_Rel_dor_fam.
CC Pfam: PF01833; TIG: 1.
CC SMART: SM00429; IPT: 1.
CC PROSITE: PS01204; REL_1; FALSE_NEG.
CC DR PROSITE: PS50254; REL_2; 1.
CC KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC Repeat; Phosphorylation.
CC FT DOMAIN 62 69 POLY-PRO.
CC FT DOMAIN 114 119 CALCINEURIN-BINDING.
CC FT REPEAT 213 293 2 APPROXIMATE SP REPEATS.
CC FT REPEAT 213 229 SP 1.
CC FT REPEAT 277 293 SP 2 (APPROXIMATE).
CC FT DOMAIN 297 304 POLY-PRO.
CC FT DOMAIN 268 270 NUCLEAR LOCALIZATION SIGNAL.
CC FT DOMAIN 430 437 DNA-BINDING.
CC FT DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.
CC SQ SEQUENCE 902 AA: 95472 MW: 855F15F7647A7C6 CRC64;

Query Match 26.6%; Score 122; DB 1; Length 902;
Best Local Similarity 40.0%; Pred. No. 2.6e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 6; Gaps 3.

Oy 2 SPAPPFQYCVETD-IPLKTRKTSDDAAILPGKLEICSDDCGNSLSPSEFTSVDDGLSGQ 60
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 SGGPFDYVGAPAESIPQKTRTSSDQAVALPRSEPPASCN-GKLPLGAESVAPPGGS- 368
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 61 YPLKRDSSGDQFLSVSPSPTWSKPK 85
   : : : | : | | | | | | | :
Db 369 ---RKEVAGMDYLAVPSPLMSKAR 390
   : : : | : | | | | | | | :

RESULT 4
NFCL_HUMAN STANDARD: PRT; 943 AA.
AC 095644; Q12865; Q15793;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 1 (NFAT transcription
DE complex cytosolic component) (NF-ATc1) (NF-ATC).
GN NFATC1 OR NFATC OR NFAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_taxonomy:9606;
NN [1]

```

RP SEQUENCE FROM N.A. (ISOFORM A-ALPHA).
 RC TISSUE=T-cell, and Peritoneal blood lymphocytes:
 RX MEDLINE=94261186; PubMed=8202141;
 RA Northrop J.P., Ho S.N., Chen L., Thomas D.J., Timmerman L.A.,
 RA Nolan G.P., Admon A., Crabtree G.R.;
 RT "NF-AT components define a family of transcription factors targeted in
 RT T-cell activation";
 RL Nature 369:497-502(1994).
 RN
 RP SEQUENCE FROM N.A. (ISOFORM B-BETA).
 RC TISSUE=B-cell;
 RX MEDLINE=96355439; PubMed=8702849;
 RA Park J., Takeuchi A., Sharma S.;
 RT "Characterization of a new isoform of the NFAT (nuclear factor of
 RT activated T cells) gene family member NFATc.";
 RL J. Biol. Chem. 271:20914-20921(1996).
 RN
 RP ERRATUM.
 RA Park J., Takeuchi A., Sharma S.;
 RL J. Biol. Chem. 271:33705-33705(1996).
 RN
 RP
 RN
 RC SEQUENCE FROM N.A. (ISOFORMS A-ALPHA; B-ALPHA AND C-BETA).
 RX MEDLINE=99170294; PubMed=10072078;
 RA Chuvpilo S., Zimmer M., Kerstan A.,
 RA Fischer C., Inashkina I., Jankevics E., Berberich-Siebel F.,
 RA Schmitt E., Serfling E.;
 RT "Alternative polyadenylation events contribute to the induction of
 RT NF-ATc in effector T cells.";
 RL Immunity 10:261-269(1999).
 RN
 RP
 RN
 RP MOTAGENESIS.
 RX MEDLINE=20119316; PubMed=10652349;
 RA Porter C.M., Havens M.A., Clipstone N.A.;
 RT "Identification of amino acid residues and protein kinases involved in
 RT the regulation of NFATc subcellular localization.";
 RL J. Biol. Chem. 275:3543-3551(2000).
 RN
 RP
 RN
 RP ALTERNATIVE SPLICING, AND CHARACTERIZATION.
 RX MEDLINE=99288090; PubMed=10358178;
 RA Chuvpilo S., Avots A., Berberich-Siebel F., Gloeckner J., Fischer C.,
 RA Kerstan A., Escher C., Inashkina I., Hlubek F., Jankevics E.,
 RA Bradletz T., Serfling E.;
 RT "Multiple NF-ATc isoforms with individual transcriptional properties
 RT are synthesized in T lymphocytes.";
 RL J. Immunol. 162:7294-7301(1999).
 RN
 RP
 RN
 RP REVIEW.
 RX MEDLINE=99189746; PubMed=10089676;
 RA Crabtree G.R.;
 RT "Genetic signals and specific outcomes: signaling through Ca2+,
 RT calcineurin, and NF-AT.";
 RL Cell 96:611-614(1999).
 CC
 CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
 CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 OR IL-4
 CC GENE TRANSCRIPTION. ALSO CONTROL GENE EXPRESSION IN EMBRYONIC
 CC CARDIAC CELLS. COULD REGULATE NOT ONLY THE ACTIVATION AND
 CC PROLIFERATION BUT ALSO THE DIFFERENTIATION AND PROGRAMMED DEATH OF
 CC T-LYMPHOCYTES AS WELL AS LYMPHOID AND NONLYMPHOID CELLS.
 CC
 CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFAT TRANSCRIPTION COMPLEX
 CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
 CC CYTOSOLIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
 CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
 CC ACTIVATING PROTEIN-1 FAMILY, MAF, GATA4 AND C/EBP-300 CAN ALSO BIND
 CC THE COMPLEX. NFAT PROTEINS BIND TO DNA AS MONOMERS.
 CC
 CC -1- SUBCELLULAR LOCATION: CYTOSOLIC FOR THE PHOSPHORYLATED FORM AND
 CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
 CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
 CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
 CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
 CC NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION.
 CC
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: A-ALPHA, A-BETA, B-
 CC ALPHA, B-BETA, C-ALPHA (SHOWN HERE) AND C-BETA. ARE PRODUCED BY


```

CC -1- DOMAIN: THE N-TERMINAL TRANSACTIVATION DOMAIN (TAD-A) BINDS TO
CC AND IS ACTIVATED BY CBP/P300. THE DEPHOSPHORYLATED FORM CONTAINS
CC TWO UNMASKED NUCLEAR LOCALIZATION SIGNALS (NLS), WHICH ALLOW
CC TRANSLLOCATION OF THE PROTEIN TO THE NUCLEUS (BY SIMILARITY).
CC -1- PPM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY
CC CALCINEURIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF069966; AAC27301.2; -
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR000451; NF_Rel_dor_fam.
DR Pfam: PF00554; RHD; 1.
DR Pfam: PF01833; TIG; 1.
DR SMART: SM00429; IPT; 1.
DR PROSITE: PS01204; REL_1; FALSE_NEG.
DR PROSITE: PS50254; REL_2; 1.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation; Repeat.
FT DOMAIN 25 32 POLY-ALA.
FT DOMAIN 110 115 CALCINEURIN-BINDING.
FT DOMAIN 118 210 TRANS-ACTIVATION DOMAIN A (TAD-A).
FT DOMAIN 195 290 3 X SP REPEATS.
FT REPEAT 195 211 1.
FT REPEAT 225 241 2.
FT REPEAT 274 290 3.
FT DOMAIN 257 259 NUCLEAR LOCALIZATION SIGNAL.
FT DOMAIN 302 313 NUCLEAR EXPORT SIGNAL.
FT DOMAIN 429 436 DNA-BINDING.
FT DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.
FT MOD_RES 109 109 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 822 AA; 88009 MW; C891D81B3644833 CRC64;

Query Match 22.2%; Score 102; DB 1; Length 822;
Best Local Similarity 32.4%; Pred. No. 0.0032;
Matches 23; Conservative 20; Mismatches 26; Indels 2; Gaps 2;

QY 16 IPLKRTKSEDAAILGRLTICSDQGNLSPSRETSDGIGSOYPLKRDSSGDFLSV 75
DB 323 VPKAKRTALDHPSPALAKVEPAEDLGATPTSDPPEPFPFQH-IRKGAFCQDYLVS 381
QY 76 PS-PFTWSPK 85
DB 382 PQHPYPMARPR 392

RESULT 6
NFCL_MOUSE STANDARD: PRT; 717 AA.
AC 088942; 070345; Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 1 (NFAT transcription
DE complex cytosolic component) (NF-ATc1) (NF-ATc).
GN NFATC1 OR NFATC OR NFAT2.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
CC [1]
RN SEQUENCE FROM N.A. (ISOFORM BETA).
RX MEDLINE=98049829; PubMed=9388475;
RA Pan S., Koyano-Nakagawa N., Tsuruta L., Amasaki Y., Yokota T.,
RA Mori S., Arai N., Arai K.-I.;
RT "Molecular cloning and functional characterization of murine cDNA

```

```

RT encoding transcription factor NFATc.";
RL Biochem. Biophys. Res. Commun. 240:314-323(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).
RC STRAIN-BAB/c; TISSUE-mast cells, and fetal liver;
RX MEDLINE=99172218; PubMed=10072529;
RA Sherman M.A., Powell D.R., Weiss D.L., Brown M.A.;
RT "NF-ATc isoforms are differentially expressed and regulated in murine
RT T and mast cells.";
RT J. Immunol. 162:2820-2828(1999).
CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 OR IL-4
CC GENE TRANSCRIPTION. ALSO CONTROL GENE EXPRESSION IN EMBRYONIC
CC CARDIAC CELLS. COULD REGULATE NOT ONLY THE ACTIVATION AND
CC PROLIFERATION BUT ALSO THE DIFFERENTIATION AND PROGRAMMED DEATH OF
CC T-LYMPHOCYTES AS WELL AS LYMPHOID AND NONLYMPHOID CELLS (BY
CC SIMILARITY).
CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
CC CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
CC ACTIVATING PROTEIN-1 FAMILY, MAF, GATA4 AND CBP/300 CAN ALSO BIND
CC THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND
CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC NFATC PLAY A KEY ROLE IN THE GENE TRANSCRIPTION (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA (SHOWN HERE) AND BETA; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. THERE MIGHT BE AN ADDITIONAL
CC ISOFORM PRODUCED BY ALTERNATIVE INITIATION AT MET-37 OF ISOFORM
CC ALPHA.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN, LUNG, SKELETAL MUSCLE,
CC THYMUS AND SKIN. WEAKLY EXPRESSED IN HEART, BRAIN, LIVER AND
CC KIDNEY. NOT EXPRESSED IN TESTIS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS AT E7 DAY AND INCREASES
CC UNTIL E17 DAY. STRONGLY EXPRESSED IN THYMUS, LUNG AND
CC SUBMANDIBULAR GLAND AND WEAKLY IN SKELETAL MUSCLE AND HEART.
CC -1- INDUCTION: INDUCIBLY EXPRESSED IN T LYMPHOCYTES UPON ACTIVATION OF
CC THE T-CELL RECEPTOR (TCR) COMPLEX. INDUCED AFTER ADDITION OF
CC PHORBOL 12-MYRISTATE 13-ACETATE (PMA).
CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).
CC -1- DOMAIN: THE N-TERMINAL TRANSACTIVATION DOMAIN (TAD-A) BINDS TO
CC AND IS ACTIVATED BY CBP/P300. THE DEPHOSPHORYLATED FORM CONTAINS
CC TWO UNMASKED NUCLEAR LOCALIZATION SIGNALS (NLS), WHICH ALLOW
CC TRANSLLOCATION OF THE PROTEIN TO THE NUCLEUS (BY SIMILARITY).
CC -1- PPM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY
CC CALCINEURIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF049606; AAC05505.1; -
DR EMBL: AF087434; AAC36725.1; -
DR MGD: MGI:102469; Nfatc1.
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR000451; NF_Rel_dor_fam.
DR Pfam: PF00554; RHD; 1.
DR Pfam: PF01833; TIG; 1.
DR SMART: SM00429; IPT; 1.
DR PROSITE: PS01204; REL_1; FALSE_NEG.
DR PROSITE: PS50254; REL_2; 1.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Alternative splicing; Phosphorylation; Repeat.

```



```
CC FT FT DOMAIN 120 125 CALCIUM-BINDING.
FT FT DOMAIN 128 220 TRANS-ACTIVATION DOMAIN A (TAD-A).
DB DB 333 VPINSRKRLALEHAPSVALKEVPEAGEDLTGTPPT-----SDFPPEETTFQHLRKGA 382
OY OY 68 SGDOFLSYVP-SPTWSSKP 85
DB DB 383 FCEGYLSYPQASQWAKPK 401
RESULT 7
DMP1_HUMAN STANDARD; PRT; 513 AA.
ID DMP1_HUMAN Q13J16; Q43Z65;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dentin matrix acidic phosphoprotein 1 precursor (dentin matrix protein-1) (DMP-1).
GN DMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Euarchonta; Primates; Catarrhini; Hominoidea; Hominidae; Homo. NCBI_TaxId=9606; [1]
RN RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Molar;
RX MEDLINE=97321043; PubMed=9177774;
RA Hirsek K.L., Simmons D., Feng J., Apelin H., Dixon M.J., McDougall M.; RT "Elucidation of the sequence and the genomic organization of the human dentin matrix acidic phosphoprotein 1 (DMP1) gene: exclusion of the locus from a causative role in the pathogenesis of dentinogenesis imperfecta type II."; RT Genomics 42:38-45(1997). RL LN NM [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
PP McDougall M., Juan X., Simmons D., Feng J.; RA Submitted (Jul-1996) to the EMBL/Genbank/DDBJ databases. RN RN [3]
RP SEQUENCE OF 462-513 FROM N.A. AND GENE MAPPING.
RA MEDLINE=96163890; PubMed=8586437;
RA Apelin H.M., Hirsek K.L., Crosby A.H., Dixon M.J.; RT "Mapping of the human dentin matrix acidic phosphoprotein gene (DMP1) to the dentinogenesis imperfecta type II critical region at chromosome 4q21."; RT Genomics 30:347-349(1995). CC -I- FUNCTION: COULD BE INVOLVED IN THE INDUCTION OF MINERALIZATION OF EXTRACELLULAR MATRIX. CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING. CC -I- TISSUE SPECIFICITY: EXPRESSED IN TOOTH PARTICULARLY IN ODONTOBLAST, AMELOBLAST AND CEMENTOBALST. CC This SWISS-PROT entry is copyright It is produced through a collaboration
```

[illegible]

```

CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; E47/PAN-1 AND E12/PAN-2 (SHOWN
CC HERE) ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- PFM: PHOSPHORYLATED FOLLOWING NGF STIMULATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
DR PIR: A37953; A37953.
DR PIR: B37953; B37953.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH_1.
DR SMART: SM00353; HLH_1.
DR PROSITE: PS00038; HLH_1; 1.
DR PROSITE: PS50088; HLH_2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW Alternative splicing; Phosphorylation.
FT DOMAIN 171 177 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DNAS_BIND 385 420 LEUCINE-ZIPPER (POTENTIAL).
FT DOMAIN 542 556 BASIC DOMAIN.
FT DOMAIN 557 600 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT VARSPLIC 528 596 DEDDLPEOKAEKEREKRRVANNARELRVINDAEFKELG
RMCOLHLSSEKQPTKLIIHOAVILS -> STEVLSLE
EKDLADRRRRMANNAREERVARDIENAREIRGICOLHLKS
DKAQTKLIILOQAVVILG (IN ISOFORM E47).
SQ SEQUENCE 649 AA: 67322 MM: 3C35F296316EB34D CRC64:

Query Match 15.3%; Score 70; DB 1; Length 649;
Best Local Similarity 26.1%; Pred. No. 6.5;
Matches 30; Conservative 6; Mismatches 17; Indels 62; Gaps 6;

OY 27 QAAIPLPKETISDDGNGISPS-----RETVDGSGSO----- 60
DB 123 QAGFLPGEGLGSLSS--PGPLSPSGVKSQGYPSYSPRRRADSLDPTQSKRVKVP 180
OY 61 -----YPLKDKSSGDOF-----LSVSPF-----TWKSP 84
DB 181 LPSSVYP---SSSGSYGRDAAPSAKTPGSAYPSPTFYVADGSLHPSAELMSP 232

RESULT 9
ALIA_STRPN STANDARD: PRT; 652 AA.
AC Q51933.
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oligopeptide-binding protein alib precursor.
GN ALIB OR SP1527.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R800;
RX MEDLINE=94328326; PubMed=8051706;
RA Allouing G., de Philip P., Claverys J.-P.;
RT "Three highly homologous membrane-bound lipoproteins participate in
RT oligopeptide transport by the Ami system of the Gram-positive
RT Streptococcus pneumoniae".
RL J. Mol. Biol. 241:44-58(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RA "Complete genome sequence of a virulent isolate of Streptococcus

```

```

RT pneumoniae.";
RL Science 293:498-506(2001).
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR OLIGOPEPTIDES. PROBABLY AN OLIGOPEPTIDE BINDING PROTEIN.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z16082; CA78896.1; -.
DR EMBL: AE007448; AA075616.1; -.
DR TIGR: SP1527; -.
DR InterPro: IPR000914; SBP_bac.5.
DR Pfam: PF00496; SBP_bac.5; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
DR PROSITE: PS01040; SBP_BACTERIAL; 1.
KW Peptide transport; Transport; Membrane; Lipoprotein; Signal;
KW Complete proteome.
FT SIGNAL 1 24 PROBABLE.
FT CHAIN 25 652 OLIGOPEPTIDE-BINDING PROTEIN ALIB.
FT LIPID 25 25 N-ACYL DIGLYCERIDE (PROBABLE).
FT CONFLICT 55 55 A -> R (IN REF. 1).
FT CONFLICT 79 80 SL -> HI (IN REF. 1).
FT CONFLICT 123 124 LQ -> FE (IN REF. 1).
FT CONFLICT 501 501 G -> E (IN REF. 1).
SQ SEQUENCE 652 AA: 72562 MM: 169B67FD78CF0CF CRC64:

Query Match 15.3%; Score 70; DB 1; Length 652;
Best Local Similarity 30.3%; Pred. No. 6.6;
Matches 27; Conservative 13; Mismatches 31; Indels 18; Gaps 5;

OY 2 SPAPFPQCVETDIPLKRTSEDOAA-----ILGKELICSDGNGISPSR---T 51
DB 29 STASKTYNIVYSSD--PSSINYLAEKNAAATSDIVANLVDELLE--NDYGNITIPSLAEDWT 85
OY 52 SYVDGLSGQYPLKRD-----SSGDOFLSY 75
DB 86 VSQDGLITTYTKRKAKAKMTSGEEYAVY 114

RESULT 10
ALIA_STRPN STANDARD: PRT; 660 AA.
AC P35592; 054782; 054620; 052228;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oligopeptide-binding protein alia precursor (Exported protein 1).
GN ALIA OR EXP1 OR PIPA OR SP0366.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R800;
RX MEDLINE=94328326; PubMed=8051706;
RA Allouing G., de Philip P., Claverys J.-P.;
RT "Three highly homologous membrane-bound lipoproteins participate in
RT oligopeptide transport by the Ami system of the Gram-positive
RT Streptococcus pneumoniae".
RL J. Mol. Biol. 241:44-58(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11906, SP-496, SP-VA92, and SP-VA96;

```

RX MEDLINE=99125733; PubMed=9466257;
 RA Coffey T.J., Enright M.C., Daniels M., Morona J.K., Morona R.,
 RA Hryniewicz M., Paton J.C., Spratt B.G.;
 RT "Recombinational exchanges at the capsular polysaccharide
 RT biosynthetic locus lead to frequent serotype changes among natural
 RT isolates of Streptococcus pneumoniae.";
 RL Mol. Microbiol. 27:73-83(1998).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN-TIGR4:
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khoult H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiolini S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 RN (4)
 RP SEQUENCE OF 18-660 FROM N.A.
 RC STRAIN-R6X:
 RX MEDLINE=95020610; PubMed=7523829;
 RA Pearce B., Naughton A.M., Masure H.R.;
 RT "Peptide permeases modulate transformation in Streptococcus
 RT pneumoniae.";
 RL Mol. Microbiol. 12:881-892(1994).
 RN (5)
 RP SEQUENCE OF 347-509 FROM N.A.
 RC STRAIN-R6X:
 RX MEDLINE=95020625; PubMed=7534910;
 RA Pearce B.J., Yin Y.B., Masure H.R.;
 RT "Genetic identification of exported proteins in Streptococcus
 RT pneumoniae.";
 RL Microbiol. 9:1037-1050(1993).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR OLIGOPEPTIDES; PROBABLY AN OLIGOPEPTIDE BINDING PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC PROTEIN FAMILY 5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z35135; CAB4507.1; -;
 DR EMBL: AE007348; AAK74534.1; ALT_INIT.
 DR EMBL: AF030359; AAC38676.1; -;
 DR EMBL: AF030360; AAC38681.1; -;
 DR EMBL: AF030361; AAC38686.1; -;
 DR EMBL: AF030364; AAC38703.1; -;
 DR EMBL: L20556; AAA26952.1; -;
 DR TIGR: SP0366; -;
 DR InterPro: IPR000914; SBP_bac_5.
 DR Pfam: PF00496; SBP_bac_5; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR PROSITE: PS01040; SBP_BACTERIAL; 5; 1.
 KW Peptide transport; Transport; Membrane; Lipoprotein; Signal;
 KW Complete proteome.
 FT SIGNAL 1 22 PROBABLE.
 FT CHAIN 23 660 OLIGOPEPTIDE-BINDING PROTEIN ALTA.
 FT LIPID 23 23 N-ACYL DIGLYCERIDE (PROBABLE).
 FT VARIANT 7 7 F -> L (IN STRAIN R800).
 FT VARIANT 27 27 G -> D (IN STRAINS NCTC 11906, SP-VA92,
 FT SP-496 AND SP-VA96).

FT VARIANT 30 30 T -> A (IN STRAINS R800 AND R6X).
 FT VARIANT 166 166 E -> D (IN STRAINS NCTC 11906, SP-VA92,
 FT SP-496 AND SP-VA96).
 FT VARIANT 246 246 V -> I (IN STRAINS R800 AND R6X).
 FT VARIANT 368 368 L -> I (IN STRAINS NCTC 11906, SP-VA92,
 FT SP-496 AND SP-VA96).
 FT VARIANT 432 432 Q -> T (IN STRAINS R800, R6X, NCTC 11906,
 FT SP-VA92, SP-496 AND SP-VA96).
 FT VARIANT 559 559 T -> A (IN STRAINS NCTC 11906, SP-VA92,
 FT SP-496 AND SP-VA96).
 FT VARIANT 612 612 T -> A (IN STRAIN SP-496).
 FT CONFLICT 18 19 TT -> GV (IN REF. 4).
 FT CONFLICT 137 137 L -> P (IN REF. 4).
 FT CONFLICT 420 420 A -> R (IN REF. 1).
 SQ SEQUENCE 660 AA; 73079 MW; 702556F92EC055E8 CRC64;
 Query Match 15.3%; Score 70; DB 1; Length 660;
 Best Local Similarity 34.1%; Pred. No. 6.7;
 Matches 28; Conservative 12; Mismatches 26; Indels 16; Gaps 6;
 OY 8 FOYCVETDIP----LKTFR--TSEDQAILPKLETCSDGDNLSPS--RETSVD-DGIAG 58
 DB 36 PSYIETDPDNLNLTITAKAATANTITSNVDCLE--NDRGNFVPMADMSVSKDGLT 93
 OY 59 SOYPLKRD----SSGDFLSV 75
 DB 94 YTYTRKDAKWYTSGEVYAAV 115
 RESULT 11
 TFE2_RAT STANDARD: PRT: 649 AA.
 ID TFE2_RAT
 AC P21677; P21676; Q08440;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription factor E2-alpha (Immunoglobulin enhancer binding factor
 DE E12/E47) (Transcription factor-3) (TCF-3) (Transcription regulator
 DE Pan).
 GN TCF3 OR PAN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORMS E12 AND E47).
 RX MEDLINE=90346284; PubMed=2200736;
 RA Nelson C., Shen L.-P., Meister A., Rodor E., Rutler W.J.;
 RT "Pan: a transcriptional regulator that binds chymotrypsin, insulin,
 RT and AP-4 enhancer motifs.";
 RL Genes Dev. 4:1035-1043(1990).
 RN (2)
 RP SEQUENCE OF 427-649 FROM N.A. (ISOFORM E12).
 RX MEDLINE=92115315; PubMed=1766666;
 RA Metz R., Ziff E.;
 RT "The helix-loop-helix protein re12 and the C/EBP-related factor rNFIL-
 RT 6 bind to neighboring sites within the c-fos serum response element.";
 RL Oncogene 6:2165-2178(1991).
 CC -1- FUNCTION: HETERODIMERS BETWEEN TCF3 AND TISSUE-SPECIFIC BASIC
 CC HELIX-LOOP-HELIX (BHLH) PROTEINS PLAY MAJOR ROLES IN DETERMINING
 CC TISSUE-SPECIFIC CELL FATE DURING EMBRYOGENESIS, LIKE MUSCLE OR
 CC EARLY B-CELL DIFFERENTIATION. DIMERS BIND DNA ON E-BOX MOTIFS: 5'-
 CC CANNTG-3'. BINDS TO THE KAPPA-E2 SITE IN THE KAPPA IMMUNOGLOBULIN
 CC GENE ENHANCER (BY SIMILARITY). BINDS TO THE CONSENSUS
 CC SEQUENCE CAC/GCTGTC PRESENT, IN THE CHMORRYSIN, INSULIN, AP-4,
 CC AND SEVERAL OTHER GENE ENHANCER MOTIFS.
 CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; E47/PAN-1 AND E12/PAN-2 (SHOWN
 CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- PTM: PHOSPHORYLATED FOLLOWING NGF STIMULATION.
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

TRANSCRIPTION FACTORS.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X54549; CAA38421.1; -
 DR EMBL: X62323; CAA44199.1; -
 DR EMBL: S77532; AAB21103.1; -
 DR PIR: A35816; A35816.
 DR PIR: B35816; B35816.
 DR TRNSPAC: T00675; -
 DR TRNSPAC: T00675; -
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH; 1.
 DR SMART: SM00353; HLH; 1.
 DR PROSITE: PS00038; HLH.1; 1.
 DR PROSITE: PS00888; HLH.2; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW Alternative splicing; Phosphorylation;
 FT DOMAIN 111 177 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT 365 420 LEUCINE-ZIPPER (POTENTIAL).
 FT 542 556 BASIC DOMAIN.
 FT 557 600 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT 525 596 PDEDLDLPKQKREKERVANNARERLRDINAFK
 FT VARSPLIC 525 596 ELGRMCOHLSTKQTKLLIHQAVYILS -> STDEV
 FT SLEEKDKRERKRNARERVRVDINDIAFEELGRMCOH
 FT LKSDKAQKRLILQAVOVITG (IN ISOFORM E47).
 FT CONFLICT 167 167 L -> LA (IN REF. 1; CAA44199).
 FT CONFLICT 427 427 P -> A (IN REF. 2).
 FT CONFLICT 508 508 D -> DH (IN REF. 2).
 FT CONFLICT 576 577 SF -> NS (IN REF. 2).
 FT CONFLICT 637 637 P -> T (IN REF. 2).
 SQ SEQUENCE 649 AA; 67654 MW; 882F19EDB47D14EA CRC64;

Query Match 15.0%; Score 69; DB 1; Length 649;
 Best Local Similarity 26.1%; Pred. No. 8.4;
 Matches 30; Conservative 6; Mismatches 17; Indels 62; Gaps 6;

QY 27 QAALLPGKLEICSDGQNLSPS-----RETSDVDGLGSG----- 60
 DB 123 QAQFLPGEGLSS--GGLPSGKSSQYTFSPNPRRAADGLDQPKKRVKPPG 180
 QY 61 -----YPLKKDSSGQF-----LSVPSPF-----TWSKP 84
 DB 181 LPSSVYP---SSSGDNYSDATAYPSAKTPSSAYBSPFYVADGSLHPSAELWSP 232

RESULT 12

NIP2_HUMAN STANDARD; PRT; 314 AA.
 AC Q12982;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE BCL2/adenovirus E1B 19-kDa protein-interacting protein 2.
 GN BNP2 OR NIP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95042730; PubMed=7954800;
 RA Boyd J.M., Malstrom S., Subramanian T., Venkatesh L.K., Schaeper U.,
 RA Elangovan B., D'Sa-Bipper C., Chinadurai G.;
 RA "Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set
 RA of cellular proteins.";
 RT Cell 79:341-351(1994).
 RL

CC -!- FUNCTION: IMPLICATED IN THE SUPPRESSION OF CELL DEATH. INTERACTS
 CC WITH THE BCL-2 AND ADENOVIRUS E1B 19 KDA PROTEINS.
 CC -!- SUBCELLULAR LOCATION: LOCALIZES TO THE NUCLEAR ENVELOPE REGION AND
 CC TO OTHER CYTOPLASMIC STRUCTURES.
 CC -!- SIMILARITY: CONTAINS 1 CRAL-TRIO DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U15173; AAC00021.1; -
 DR Genew: HGNC:1083; BNP2.
 DR MIM: 603292;
 DR InterPro: IPR001251; CRAL-TRIO.
 DR SMART: SM00516; SEC14; 1.
 DR PROSITE: PS50191; CRAL-TRIO; 1.
 KW Apoptosis.
 FT DOMAIN 147 304 CRAL-TRIO.
 SQ SEQUENCE 314 AA; 36018 MW; 0A3CEC6356569DD2F CRC64;

Query Match 14.7%; Score 67.5; DB 1; Length 314;
 Best Local Similarity 31.3%; Pred. No. 5;
 Matches 31; Conservative 7; Mismatches 36; Indels 25; Gaps 7;

QY 3 PAPFFQYVEVDIPLKTKTSEDAALPGKLEICSDGQNLSPSRETSDVDG 56
 DB 16 PIPLEPDDSIEDILAIT--GPEQD---PGSLEV---NGKKVRKKLMAPDISILTPDS 65
 QY 57 LGSQYPLKKDSSGD---OFLSPVS---PFTWSK--PKP 86
 DB 66 DGSVLSDDLDDESGEIDLDGDPSPSENSFEWEDDLPRK 104

RESULT 13

VNML_DROME STANDARD; PRT; 558 AA.
 AC Q9NFP1; Q9M431;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vanin-like protein precursor.
 GN CG3648.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99432008; PubMed=10501839;
 RA Granjeaud S., Naquet P., Galland F.;
 RA "An EST description of the new Vanin gene family conserved from fly to
 RA human.";
 RT Immunogenetics 49:964-972(1999).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.Y.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glöckner G., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jamali M., Katush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jinnai B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasero P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mactel B., McInosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusseren D.J., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheelers F., Shen H.,
RA Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weisslock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: PROBABLE HYDROLASE.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
(potential).
CC -1- TISSUE SPECIFICITY: Expressed in larvae and early pupae.
CC -1- SIMILARITY: BELONGS TO THE CN HYDROLASE FAMILY. BTD/VNN SUBFAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ276261; CAB77020.1; -
DR EMBL: AE003436; AAF6129.1; ALT_SBO.
DR FlyBase: FBgn0040069; vanin-like.
DR InterPro: IPR003010; Ntase/CNhydase.
DR Pfam: PF00795; CN_hydrolase; 1.
KM Hydrolase; Signal; Glycoprotein; GPI-anchor.
FT SIGNAL 1 22
FT CHAIN 1 22
FT PROPEP 533 532
FT LIPID 532 532
FT CARBOHYD 65 65
FT CARBOHYD 103 103
FT CARBOHYD 120 120
FT CARBOHYD 128 128
FT CARBOHYD 180 180
FT CARBOHYD 354 354
FT CARBOHYD 379 379
SQ SEQUENCE 558 AA: 62341 MW: 7920621923587779 CRC64;
Query Match 14.6% Score 67; DB 1; Length 558;
Best Local Similarity 32.8% Pred. NO. 11;
Matches 22; Conservative 5; Mismatches 22; Indels 18; Gaps 3;
OY 38 CGDD-----GNSPSRSTSVDDGLASQIP-----LKKDSSDGLV-PSPF 79
DB 412 CGGSDIDCGKLLPTEGELQOSRVFTRLAIGVTYPSRREFLLFPDTLQDLSLPLEPSOF 471
OY 80 TWSKRP 86
DB 472 EWSKRP 478

RESULT 14
ID RIK3_MOUSE STANDARD: PRT; 486 AA.
AC 09QZL0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor-interacting serine/threonine protein kinase 3 (EC 2.7.1.-)
DE (RIP-like protein kinase 3) (Receptor-interacting protein 3) (RIP-3)
DE (mrip3).
GN RIKP3 OR RIP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090.
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ASP-143.
RC TISSUE=Embryo;
RX MEDLINE=99421935; PubMed=10490590;
RA Padzernik N.J., Donner D.B., Goebel M.G., Harrington M.A.;
RT "Mouse receptor interacting protein 3 does not contain a caspase-
RT recruiting or a death domain but induces apoptosis and activates NF-
RT kappaB.";
RL Mol. Cell. Biol. 19:6500-6508(1999).
CC -1- FUNCTION: Promotes apoptosis.
CC -1- SUBUNIT: Binds TRAF2 and RIPK1 and is recruited to the TNFR-1
CC signaling complex (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: Expressed in embryo and in adult spleen,
CC liver, testis, heart, brain and lung.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF178953; AAF03133.1; -
DR MGD: MGI:2154952; Ripk3.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferase; Serine/threonine-protein kinase; ATP-binding;
KM Phosphorylation; Apoptosis.
FT DOMAIN 22 292
FT NP_BIND 28 36
FT BINDING 51 51
FT ACT_SITE 143 143
FT DOMAIN 352 438
FT MUTAGEN D->N: NO AUTOPHOSPHORYLATION.
SQ SEQUENCE 486 AA: 53336 MW: DD264E69187D3436 CRC64;
Query Match 14.5% Score 66.5; DB 1; Length 486;
Best Local Similarity 24.8% Pred. NO. 11;
Matches 28; Conservative 12; Mismatches 36; Indels 37; Gaps 6;
OY 7 PFQYCVENDIPLKT-----RKTSEDQAALPGKLEICSDQ-----GNLSPSRSTSV 53
DB 325 PSQRCETMDCPRETWVSKMLDLHLEPPSGVPVCK--CPENQADTSVGPATPAR-TSS 380
OY 54 DGLGS-----QYPLKDDSGDGLSVSP-----FTWSKRP 86

DB 381 DPVAGTPOIPIHTLPFGTTPGVTETPCPHORNGDGRHGTWPWTPPNP 433

RESULT 15

HGF_MOUSE ID HGF_MOUSE STANDARD: PRT: 728 AA.

AC 008048; 064007; 061662;

DR 01-NOV-1995 (Rel. 32, Created)

DR 01-NOV-1995 (Rel. 32, Last sequence update)

DR 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hepatocyte growth factor precursor (scatter factor) (SF)

GN (Hepatopoietin A).

OS HGF.

OC Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RM [1]

RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND SEQUENCE OF 496-504.

RC TISSUE=Mammary fibroblast;

RA MEDLINE=94183257; PubMed=8135822;

RA Sasaki M., Nishio M., Sasaki T., Enami J.;

RT "Identification of mouse mammary fibroblast-derived mammary growth factor as hepatocyte growth factor."

RT Biochem. Biophys. Res. Commun. 199;772-779(1994).

RL [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=94363381; PubMed=8081873;

RA Lee C.C., Kozak C.A., Yamada K.M.;

RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter factor gene."

RT Cell Adhes. Commun. 1;101-111(1993).

RL [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=94060105; PubMed=8241272;

RA Liu Y., Michalopoulos G.K., Zarnegar R.;

RT "Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth factor."

RT Biochim. Biophys. Acta 1216;299-303(1993).

CC -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL HEPATOCYTE CELLS. SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES. IT HAS NO DETECTABLE PROTEASE ACTIVITY.

CC -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A DISULFIDE BOND.

CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a short form; are produced by alternative splicing.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.

CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: D10212; BA01064.1; -

DR EMBL: D10213; BA01065.1; -

DR EMBL: S71816; AAB31855.1; -

DR EMBL: X72307; CAAS1054.1; ALT_INIT.

DR HSSP: P14210; IBHT.

DR MEROPS: S01.982; -

DR MGD: MGT.96079; HGT.

DR InterPro: IPR001314; Chymotrypsin.

DR InterPro: IPR000001; Kringle.

DR InterPro: IPR003014; PAN.

DR InterPro: IPR003609; Pan_app.

DR InterPro: IPR001254; Ser_protease_Try.

DR Pfam: PF00024; PAN; 1.

DR Pfam: PF00051; kringle; 4.

DR Pfam: PF00089; trypsin; 1.

DR PRINTS: PR00722; CHYMOTRYPSIN.

DR PRINTS: PR00018; KRINGLE.

DR PRODOM: PD000395; Kringle; 4.

DR SMART: SM00130; KR; 4.

DR SMART: SM00473; PAN_AP; 1.

DR SMART: SM00020; TRY-Spc; 1.

DR PROSITE: PS00021; KRINGLE 1; 4.

DR PROSITE: PS00070; KRINGLE 2; 4.

DR PROSITE: PS0240; TRYPSIN_DOM; 1.

KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;

KW Repeat; Signal; Alternative splicing.

FT SIGNAL 1 32

FT CHAIN 33 495

FT CHAIN 496 728

FT MOD_RES 33 33

FT DOMAIN 33 128

FT DOMAIN 129 207

FT DOMAIN 212 289

FT DOMAIN 306 384

FT DOMAIN 392 470

FT DOMAIN 496 728

FT DISULFD 71 97

FT DISULFD 75 85

FT DISULFD 488 607

FT CARBOHYD 295 295

FT CARBOHYD 403 403

FT CARBOHYD 569 569

FT CARBOHYD 656 656

FT VARSPLC 163 167

FT CONFLICT 344 344

FT CONFLICT 479 479

FT CONFLICT 564 564

SO SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CnC64;

Query Match 14.4%; Score 66; DB 1; Length 728;

Best Local Similarity 23.5%; Pred. No. 20;

Matches 19; Conservative 12; Mismatches 20; Indels 30; Gaps 3;

OY 7 PROYCV-----ENDIPKTRKTSDDAAILPKGLICSDGQGLSPSRFISYVDG 56

DB 280 PWEYCAITCAHSAVAVETDVPMETTE-----CIGQGEGYRGTSTNTWNG 324

OY 57 L-----GSQYPLKRDSSGDOF 72

DB 325 IPCQRWDSQYPRKHDTTPENF 345

Search completed: April 11, 2003, 14:22:42

Job time : 16 secs


```

DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE 3110041H08RIK protein (Fragment).
GN NFATC4 OR 3110041H08RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Aachji J., Fukuda S.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RL EMBL: AK014164; BAB29185.1;
DR MGD: MGI:1920431; Nfatc4.
FT NON_TER 1
SQ SEQUENCE 226 AA; 23584 MW; 882F2BBD4FEDB595 CRC64;

Query Match 28.1%; Score 129; DB 11; Length 226;
Best Local Similarity 34.1%; Pred. No. 9e-07;
Matches 29; Conservative 18; Mismatches 32; Indels 6; Gaps 2;

QY 2 SPAPPEFCVETD-IPKTRKTSDDAAILPGKLEICSDGDNLSPSRETSVDGLGSG 60
DB 138 SPGPFYVAGPPTESIPQKTRRTSSQAVALPR-----SEPPSCNGKLPSTGDSVAP 192
QY 61 YPLKDDSGDQFLSVSPPTWSKPK 85
DB 193 GALRKEVAGMDYLAVPSPLAWSKAR 217

RESULT 3
Q9EP91 PRELIMINARY; PRT; 901 AA.
AC Q9EP91:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Transcription complex subunit NF-ATC4.
GN NFATC4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV/EV;
RA Graef I.A., Crabtree G.R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C129SV/EV;
RA Graef I.A., Crabtree G.R.;

```

```

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF309389; AAG39446.1; -.
DR EMBL: AF309388; AAG39446.1; JOINED.
DR EMBL: AF283284; AAF98174.1; -.
DR MGD: MGI:1920431; Nfatc4.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR000451; NF_Rel_dor_fam.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF01833; TIG; 1.
DR PRINTS: PR01217; PRICHEXTENSN.
DR SMART: SM00429; IPT; 1.
DR PROSITE: PS00455; AMP_BINDING; UNKNOWN_1.
DR PROSITE: PS50254; REL_2; 1.
SQ SEQUENCE 901 AA; 95820 MW; 7F66A93E76FDB773 CRC64;

Query Match 27.5%; Score 126; DB 11; Length 901;
Best Local Similarity 32.9%; Pred. No. 1.1e-05;
Matches 28; Conservative 19; Mismatches 32; Indels 6; Gaps 2;

QY 2 SPAPPEFCVETD-IPKTRKTSDDAAILPGKLEICSDGDNLSPSRETSVDGLGSG 60
DB 311 SPGPFYVAGPPTESIPQKTRRTSSQAVALPR-----SEPPSCNGKLPSTGDSVAP 365
QY 61 YPLKDDSGDQFLSVSPPTWSKPK 85
DB 366 GALRKEVAGMDYLAVPSPLAWSKAR 390

RESULT 4
Q96H68 PRELIMINARY; PRT; 845 AA.
AC Q96H68:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Unknown (Protein for IMAGE:4109469) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008857; AAH08857.1; -.
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR000451; NF_Rel_dor_fam.
DR Pfam: PF01833; TIG; 1.
DR PROSITE: PS50254; REL_2; 1.
FT NON_TER 1
SQ SEQUENCE 845 AA; 89682 MW; 257F079E2DEA5F16 CRC64;

Query Match 26.6%; Score 122; DB 4; Length 845;
Best Local Similarity 40.0%; Pred. No. 2.8e-05;
Matches 34; Conservative 11; Mismatches 34; Indels 6; Gaps 3;

QY 2 SPAPPEFCVETD-IPKTRKTSDDAAILPGKLEICSDGDNLSPSRETSVDGLGSG 60
DB 254 SPGPFYVAGPPTESIPQKTRRTSSQAVALPRSEPPASCN-GKLPGLAEESVAPPGS- 311
QY 61 YPLKDDSGDQFLSVSPPTWSKPK 85
DB 312 ---RKEVAGMDYLAVPSPLAWSKAR 333

RESULT 5
Q9DB06 PRELIMINARY; PRT; 703 AA.
AC Q9DB06:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)

```


DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Nuclear factor of activated T-cells, cytoplasmic 1.
CN NPATC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai Y., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi S., Fukuda S.,
RA Akawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Knehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Fletcher C., Fujita M., Gariboldi M.,
RA Giustolisi S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Nombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
DR EMBL: AK004810; BAB3581.1; -
DR MGD: MGI:102469; Nfatac1.
DR InterPro: IPRO002909; IPT_TIG.
DR InterPro: IPRO00451; NF_Rel_dor_fam.
DR Pfam: PF00554; RHD: 1.
DR Pfam: PF01833; TIG: 1.
DR SMART: SM00429; IPT: 1.
DR PROSITE: PS50254; REL_2: 1.
SQ SEQUENCE 703 AA; 76691 MW; D737C9B86E280C7D CRC64;

Query Match 21.0%; Score 96.5; DB 11; Length 703;
Best Local Similarity 30.4%; Pred. No. 0.018;
Matches 24; Conservative 17; Mismatches 19; Indels 19; Gaps

OY 16 IPLKTRKTSEDOAMILPKLEICSDQGNLSRFTSVDDGLGSOYP-----LKXDS 67
DB 319 VPIKSRTALEHAHPASVALKVBPAGBDLGTPPT-----SDFPPEETFGHLRKKA 368
DY 68 SGDOFLSP-SPTFSKRPK 65
DB 369 FCEOYLSPQASYNWAKPK 387

RESULT 6
O3JHJD3 PRELIMINARY: PRF: 717 AA.

AC O3JHJD3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Nuclear factor of activated T cells C.
CN NPATC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98049829; PubMed=9388475;
RA Pan S., Koyano-Nakagawa N., Tsuruta L., Amasaki Y., Yokota T.,
RA Motl S., Atal N., Atal K.;
RT "Molecular cloning and functional characterization of murine CDNA

[illegible]

```

RESULT 8
065389 PRELIMINARY: PRT: 329 AA.
AC 065389;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE F12fl.23 protein (Hypothetical 36.7 kDa protein).
GN F12fl.23.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Vysotskaia V.S., Osborne B.I., Schwartz J.R., Toriumi M., Yu G.,
RA Kwan A., Oji O., Liu S., Buehler E., Conway A.B., Conway A.R.,
RA Dewar K., Feng J., Kim C., Kurtz D., Palm C.J., Li Y., Shin P.,
RA Sun H., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Kalin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene F12fl.23 (GI:3157936).";
RT Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Kalin-Neumann G., Kawai J., Kim C., Koesema E., Lam B.,
RA Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shin P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F12fl.23 (GI:3157936).";
RT Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC002131; AAC17619.1;
DR EMBL: AF360349; AAK28646.1;
DR EMBL: AY051085; AAK93762.1;
KW Hypothetical protein.
SQ SEQUENCE 329 AA; 36742 MW; 4A87A14ECA5A1CC CRC64;

Query Match 17.8%; Score 81.5; DB 10; Length 329;
Best Local Similarity 25.3%; Pred. No. 0.37;
Matches 20; Conservative 10; Mismatches 26; Indels 23; Gaps 1;

QY 27 QAAIIPGKLEISDQGNLSPRETSDVGLGS-----QYPL 63
DB 20 QSSITSSQSNLCRSSCGNIPINYPISIDGCGSPYRMLICSDNDTKLELTSPGKRYV 79
OY 64 KRDSSGDOFLVSPPTWS 82
DB 80 KSISYSDPHLTVSDPFMN 98

RESULT 9
09FJL6 PRELIMINARY: PRT: 1065 AA.
AC 09FJL6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)

```

```

DE: Gb|AAD25584.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL: AB013396; BAB08870.1;
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR004274; NIF.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF03031; NIF; 1.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS50172; BRCT; 1.
SQ SEQUENCE 1065 AA; 120932 MW; 689B6DFBF486CF0 CRC64;

Query Match 17.2%; Score 79; DB 10; Length 1065;
Best Local Similarity 33.9%; Pred. No. 2.9;
Matches 20; Conservative 11; Mismatches 26; Indels 2; Gaps 1;

QY 1 LSPAPPFOYCVETDIPKTRKTSDDAAILPGKLEICSDQGNLSPRETSDVGLGS 59
DB 617 LSPPTNKFMSVAVSDSPVHSSSSDDLAFDLAFLDASD--ASSGSPSEEEAEDVES 673

RESULT 10
09I7N6 PRELIMINARY: PRT: 1257 AA.
AC 09I7N6; Q9TVE8; Q9TVE9; Q9Y195; Q9UA38; Q9U971;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE DNA-repair protein complementing XP-G cells homolog (Xeroderma
DE pigmentosum group G complementing protein homolog) (Nucleo-
DE 201) (Nucleotide excision repair protein mus201) (DNA repair
DE endonuclease).
DE MUS201 OR XPG OR CG10890.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=EMBRYO.
RX MEDLINE=99326510; PubMed=10395909;
RA Houle J.-F., Friedberg E.C.;
RT "The Drosophila ortholog of the human XPG gene.";
RL Gene 234:353-360(1999).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=EMBRYO.
RX MEDLINE=20274058; PubMed=10812334;
RA Sekelsky J.J., Hollis K.J., Eimerl A.I., Burtis R.C., Hawley R.S.;
RT "Nucleotide excision repair endonuclease genes in Drosophila
RT melanogaster.";
RL Mutat. Res. 459:219-228(2000).
RN [3]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RA Calleja F.M.G.R., Nivard M.J.M., Vogel E.W., Ecken J.C.J.;
RT "Genetic effects in the Drosophila mutant mus201[D1] are caused by
RT alterations in the Drosophila homolog of the nucleotide excision
RT repair gene encoding XPG.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: Apr11 11, 2003, 14:21:47 ; Search time 14 Seconds
(without alignments)
180.741 Million cell updates/sec

Title: US-09-550-115a-2

Perfect score: 459

Sequence: 1 LSPAPFPQYCVETDIPKTRKTS...SSGQFLSVSPFTWSKPKP 86

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	430	93.7	708	1 US-08-396-479B-8	Sequence 8, Appl
2	430	93.7	708	1 US-08-818-823-8	Sequence 8, Appl
3	430	93.7	739	1 US-08-396-479B-10	Sequence 10, Appl
4	430	93.7	739	1 US-08-818-823-10	Sequence 10, Appl
5	430	93.7	1068	1 US-08-396-479B-12	Sequence 12, Appl
6	430	93.7	1068	1 US-08-818-823-12	Sequence 12, Appl
7	430	93.7	1075	5 PCT-US94-07297-41	Sequence 41, Appl
8	122	26.6	902	1 US-08-396-479B-6	Sequence 6, Appl
9	122	26.6	902	1 US-08-818-823-6	Sequence 6, Appl
10	114	24.8	716	1 US-08-396-479B-4	Sequence 4, Appl
11	114	24.8	716	1 US-08-818-823-4	Sequence 4, Appl
12	114	24.8	716	3 US-09-037-190-38	Sequence 38, Appl
13	114	24.8	716	3 US-09-037-190-46	Sequence 46, Appl
14	114	24.8	716	3 US-09-037-192-38	Sequence 38, Appl
15	114	24.8	716	3 US-09-037-192-46	Sequence 46, Appl
16	114	24.8	716	4 US-09-037-143-38	Sequence 38, Appl
17	114	24.8	716	4 US-09-037-143-46	Sequence 46, Appl
18	114	24.8	716	4 US-09-049-691-38	Sequence 38, Appl
19	114	24.8	716	4 US-09-049-691-46	Sequence 46, Appl
20	114	24.8	716	4 US-08-260-174-38	Sequence 38, Appl
21	114	24.8	716	4 US-08-260-174-46	Sequence 46, Appl
22	114	24.8	716	4 US-09-338-128A-38	Sequence 38, Appl
23	114	24.8	716	4 US-09-338-128A-46	Sequence 46, Appl
24	114	24.8	716	4 US-09-233-346-38	Sequence 38, Appl
25	114	24.8	716	4 US-09-233-346-46	Sequence 46, Appl
26	114	24.8	716	4 US-09-037-192-38	Sequence 38, Appl
27	114	24.8	716	4 US-09-037-192-46	Sequence 46, Appl

28	114	24.8	716	5 PCT-US94-07297-37	Sequence 37, Appl
29	114	24.8	761	2 US-08-124-981A-2	Sequence 2, Appl
30	70	15.3	643	2 US-08-245-511-47	Sequence 47, Appl
31	70	15.3	643	2 US-08-600-993A-47	Sequence 47, Appl
32	67.5	14.7	314	2 US-08-408-095-23	Sequence 23, Appl
33	64	13.9	422	4 US-08-790-186A-4	Sequence 4, Appl
34	63	13.7	398	3 US-09-045-632-31	Sequence 31, Appl
35	63	13.7	441	3 US-09-045-632-37	Sequence 37, Appl
36	63	13.7	498	3 US-09-045-632-30	Sequence 30, Appl
37	63	13.7	541	3 US-09-045-632-36	Sequence 36, Appl
38	63	13.7	599	3 US-09-045-632-28	Sequence 28, Appl
39	63	13.7	642	3 US-09-045-632-35	Sequence 35, Appl
40	63	13.7	818	3 US-09-045-632-25	Sequence 25, Appl
41	63	13.7	861	3 US-09-045-632-34	Sequence 34, Appl
42	63	13.7	918	3 US-09-045-632-31	Sequence 31, Appl
43	63	13.7	961	3 US-09-045-632-33	Sequence 33, Appl
44	63	13.7	1018	3 US-09-045-632-16	Sequence 16, Appl
45	63	13.7	1061	3 US-09-045-632-32	Sequence 32, Appl

ALIGNMENTS

```
RESULT 1
US-08-396-479B-8
Sequence 8, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESS: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 708 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-396-479B-8
Query Match 93.7% Score 430: DB 1: Length 708;
Best Local Similarity 91.9% Pred. No. 2e-47; 3: Indels 0: Gaps 0:
Matches 79: Conservative 4: Mismatches
```

OY 1 LSPAPFPQYCVETDIPKTRKTS...SSGQFLSVSPFTWSKPKP 60
DB 321 LSPAPFPQYCVETDIPKTRKTS...SSGQFLSVSPFTWSKPKP 380
OY 61 YPKKDSGQFLSVSPFTWSKPKP 86
|||||

DB 381 YPLKDCGDFLSVSPPTWSKPKP 406

RESULT 2

US-08-818-823-8

Sequence 8, Application US/08818823

Patent No. 5708158

GENERAL INFORMATION:

APPLICANT: HOEY, Timothy

TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/818, 823

FILING DATE: 14-MAR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/396,479

FILING DATE: 02-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59450-1/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 210 277299

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 708 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-818-823-8

Query Match

Best Local Similarity 93.7%; Score 430; DB 1; Length 708;

Matches 79; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSPAPFPQYCVETDIPKTRKTSDDAAILPGKLEICSDDGNI:SPSRSTVDDGLGSO 60

DB 321 LCPAVFPQYCVETDIPKTRKTSDDAAILPGKLEICSDDGSI:SPARSTIDGGLGSO 380

QY 61 YPLKDCGDFLSVSPPTWSKPKP 86

DB 381 YPLKDCGDFLSVSPPTWSKPKP 406

RESULT 3

US-08-396-479B-10

Sequence 10, Application US/08396479B

Patent No. 5612455

GENERAL INFORMATION:

APPLICANT: HOEY, Timothy

TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/396,479B

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59450-1/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 210 277299

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 739 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-396-479B-10

Query Match

Best Local Similarity 93.7%; Score 430; DB 1; Length 739;

Matches 79; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSPAPFPQYCVETDIPKTRKTSDDAAILPGKLEICSDDGNI:SPSRSTVDDGLGSO 60

DB 321 LCPAVFPQYCVETDIPKTRKTSDDAAILPGKLEICSDDGSI:SPARSTIDGGLGSO 380

QY 61 YPLKDCGDFLSVSPPTWSKPKP 86

DB 381 YPLKDCGDFLSVSPPTWSKPKP 406

RESULT 4

US-08-818-823-10

Sequence 10, Application US/08818823

Patent No. 5708158

GENERAL INFORMATION:

APPLICANT: HOEY, Timothy

TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/818, 823

FILING DATE: 14-MAR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/396,479

FILING DATE: 02-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59450-1/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771


```

: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/396,479B
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59450-1/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELEX: 210 277299
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 902 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-396-479B-6

Query Match 26.6%; Score 122; DB 1; Length 902;
Best Local Similarity 40.0%; Pred. No.3,5e-07;
Matches 34; Conservative 11; Mismatches 34; Indels 6; Gaps 3

OY 2 SPAPEPDYCYETD-IPLKTRKTSDDQAILPGKLEICSDDGNGLSPSRFISVDGLGSG 60
: |||: : |||||: ||| : ||| : |||
Db 311 SPGPDYVAGAPPAEISIPKTRTSSDOAVLPRSEEPASCN-GKILPGAESVAPPGGS- 368
: ||: : ||: ||||| ||| :

OY 61 YPLKDDSGDGFVSPSPFTWSPK 85
: ||: : ||: ||||| ||| :

Db 369 ---RKEVAGMDYLAVPSPPLANSKAR 330

RESULT 9
US-08-818-823-6
: Sequence 6, Application US/08818823
: Patent No. 5708158
: GENERAL INFORMATION:
: APPLICANT: HOEY, Timothy
: TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/818,823
: FILING DATE: 14-MAR-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/396,479
: FILING DATE: 02-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59450-1/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELEX: 210 277299
: SEQUENCE CHARACTERISTICS:
: LENGTH: 902 amino acids
: TYPE: amino acid

```

```

;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-818-823-6

```

Query Match	26.68;	Score 122;	DB 1;	Length 902;
Best Local Similarity	-40.08;	Pred. NO. 3.5e-07;		
Matches 34;	Conservative 11;	Mismatches 34;	Indels 6;	Gaps 3;

Oy 2 SPAPPPQYCVETD-IPLKTRTSEDQAAILPGLALEICSDQGNSPSRETSVDGIGSQ 60
 ||| : ||||| : | : | : |
Db 311 SPCPFDYGVAPAEISLPQKTRRTSSSEQVALPSEEPASCN-GKLPLGAEESSVAPPGGS - 368

```
OY      61 YPLKKDSSGDDFLSVSPPTWSKPK 85  
          :|::||| |  
Db     369 ---RKEVAGMDYLAVPSPLAWSKAR 390
```

RESULT 10
US-08-396-479B-4
; Sequence 4, Application US/08396479B
; Patent No. 5612455

1 GENERAL INFORMATION:
 2 APPLICANT: HOEY, Timothy
 3 TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
 4
 5 NUMBER OF SEQUENCES: 18
 6
 7 CORRESPONDENCE ADDRESS:
 8
 9 ADDRESSEE: FLEHN, HOBACH, TEST, ALBRITTON & HERBERT
 10 STREET: 4 Embarcadero Center, Suite 3400
 11 CITY: San Francisco
 12 STATE: California
 13 COUNTRY: USA

```

:      COMPUTER READABLE FORM:
:      MEDIUM TYPE: Floppy disk
:      .COMPUTER: IBM PC compatible
:      OPERATING SYSTEM: PC-DOS/MS-DOS
:      SOFTWARE: PatentIn Release #1.0, Version #1.30
:      CURRENT APPLICATION DATA:

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771

```

: INFORMATION FOR SEQ ID NO: 4
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 716 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
US-08-396-479B-4

```

Query Match	24.8%	Score 114;	DB 1;	Length 716;
Best Local Similarity	34.7%	Pred. No. 2.8e-06;		
Matches	25;	Conservative 19;	Mismatches 26;	Indels 2;
			Gaps	2

```
Oy      16 IPLTRKTSEQAMILPGKLEICSDDOGNLSPRHETSVDDGLGSOYPLKKSSGGDFLSV 75
       ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db     331 VPVSRKRTTLEQPPSYALKVEPVGEDLGSPPPADFAPEDYSSFQH-IRKGFCDDYLAV 389
```

```

QY      76 PS-PETWSKPKP 86
        | : : : | | |
Db      390 PQHPYQWAKPKP 401

```

RESULT 11
US-08-818-823-4
: Sequence 4, Application US/08818823

```
; Patent No. 5708158
;
; GENERAL INFORMATION:
;
; APPLICANT: HOEY, Timothy
;
; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
;
; NUMBER OF SEQUENCES: 18
;
; COMPOUND(S) AND REAGENT(S)
```

ADDRESSEE: FLEHR, HOMBACH, TEST, ALBRITTON & HEBBER
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:

```

1 PRIOR APPLICATION DATA: 08/396,479
2 APPLICATION NUMBER: 02-MAR-1995
3 FILING DATE: 02-MAR-1995
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Osman, Richard A.
6 REGISTRATION NUMBER: 36,627
7 REFERENCE/DOCKET NUMBER: A-59450-1/RAC
8 TELECOMMUNICATION INFORMATION:

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 716 amino acids
; TYPE: amino acid
;

```

Query Match	24.88;	Score 114;	DB 1;	Length 716;
Best Local Similarity	34.78;	Pred. No. 2.8e-06;		
Matches 25; Conservative	19;	Mismatches 26;	Indels 2;	Gaps 2

```
QY 16 IPLKTRKTSIEDQAILPGKLEICSDDOGNLSPSRETSYVDGLGSOYPLKKDSSGDQFLSV 75
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 331 VPKSRKRTTLEQPSVALKVEPVGEDGSPPPADFAPEDYSSFOH-IRKGGFCDDYLVAV 389
```

QY	76	PS-PTWSKPKP	86
		: :	
Db	390	QAHRYQWAKPKP	401

RESULT 12
US-09-037-190-38
; Sequence 38, Application US/09037190

```

: GENERAL INFORMATION:
: APPLICANT: Crabtree, Gerald R.
: APPLICANT: NO. 609651stthrop, Jeffrey P.
: APPLICANT: HO, Stetlan M.
: TITLE OF INVENTION: NE-AT POLYPEPTIDES AND POLYNUCLEOTIDES
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:

```

ADDRESS: FOLEY, HONG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 14:22:26 ; Search time 14 Seconds
(without alignments)
375.550 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459

Sequence: 1 LSPAPFPFOYCVETDIPLKT.....SSGDFLSVSPFTWSPKPK 86

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB pep: *
2: /cgn2_6/ptodata/2/pubpaa/PC1_NEM_PUB pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB pep: *
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7: /cgn2_6/ptodata/2/pubpaa/PC105_PUBCOMB pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB pep: *
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep: *
11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB pep: *
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep: *
13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB pep: *
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	24.8	716	US-10-040-430-38	Sequence 38, Appl
2	114	24.8	716	US-10-040-430-46	Sequence 46, Appl
3	114	24.8	716	US-09-349-058-38	Sequence 38, Appl
4	114	24.8	716	US-09-349-058-46	Sequence 46, Appl
5	65.5	14.3	557	US-09-925-289-940	Sequence 940, App
6	65.5	14.3	557	US-09-925-289-940	Sequence 940, App
7	64.5	14.1	694	US-09-925-289-940	Sequence 940, App
8	63	13.7	404	US-09-764-864-1244	Sequence 1244, App
9	62	13.5	522	US-09-765-272-120	Sequence 120, App
10	62	13.5	538	US-09-765-272-120	Sequence 120, App
11	62	13.5	1040	US-09-765-272-118	Sequence 118, App
12	61.5	13.4	427	US-09-965-529-20	Sequence 20, Appl
13	61.5	13.4	1368	US-10-159-151-2	Sequence 2, Appl
14	61	13.3	112	US-09-216-393-83	Sequence 83, Appl
15	61	13.3	1203	US-10-027-923-5	Sequence 5, Appl
16	60.5	13.2	174	US-09-738-626-6602	Sequence 6602, App
17	60	13.1	491	US-10-227-884-34	Sequence 34, Appl
18	60	13.1	491	US-10-230-163-34	Sequence 34, Appl
19	60	13.1	491	US-10-218-631-34	Sequence 34, Appl

20	60	13.1	491	US-10-230-338-34	Sequence 34, Appl
21	60	13.1	491	US-10-230-414-34	Sequence 34, Appl
22	60	13.1	577	US-09-810-264-24	Sequence 24, Appl
23	60	13.1	1213	US-09-969-515-8	Sequence 8, Appl
24	60	13.1	1224	US-09-969-515-2	Sequence 2, Appl
25	59.5	13.0	467	US-09-912-020-381	Sequence 381, App
26	59	12.9	218	US-09-825-561A-6	Sequence 6, Appl
27	59	12.9	289	US-09-941-973-2	Sequence 2, Appl
28	59	12.9	349	US-09-809-391-371	Sequence 371, App
29	59	12.9	538	US-09-972-218A-2	Sequence 2, Appl
30	59	12.9	538	US-09-965-113-2	Sequence 2, Appl
31	59	12.9	538	US-09-923-246-115	Sequence 115, App
32	59	12.9	538	US-09-825-561A-2	Sequence 2, Appl
33	59	12.9	567	US-09-825-561A-16	Sequence 16, Appl
34	59	12.9	606	US-09-923-246-97	Sequence 97, Appl
35	59	12.9	606	US-09-825-561A-30	Sequence 30, Appl
36	59	12.9	911	US-09-291-417-92	Sequence 92, Appl
37	59	12.9	3353	US-09-888-615-64	Sequence 64, Appl
38	58.5	12.7	592	US-10-227-884-62	Sequence 62, Appl
39	58.5	12.7	592	US-10-230-163-62	Sequence 62, Appl
40	58.5	12.7	592	US-10-218-631-62	Sequence 62, Appl
41	58.5	12.7	592	US-10-230-338-62	Sequence 62, Appl
42	58.5	12.7	592	US-10-230-414-62	Sequence 62, Appl
43	58.5	12.7	720	US-09-801-368-176	Sequence 176, App
44	58	12.6	129	US-09-864-761-41836	Sequence 41836, A
45	58	12.6	181	US-09-984-271-114	Sequence 114, App

ALIGNMENTS

RESULT 1
US-10-040-430-38
Sequence 38, Application US/10040430
Publication NO. US20030049641A1
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: No. US20030049641A1throp, Jeffrey P.
APPLICANT: Ho, Stefan M.
APPLICANT: Flanagan, William M.
TITLE OF INVENTION: NP-AT POLYPEPTIDES AND POLYNUCLEOTIDES AND SCREENING
TITLE OF INVENTION: METHODS FOR IMMUNOSUPPRESSIVE AGENTS
FILE REFERENCE: APV-008.04
CURRENT APPLICATION NUMBER: US/10/040.430
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: US/09/232.346
PRIOR FILING DATE: 1999-01-15
PRIOR APPLICATION NUMBER: 08/507,032
PRIOR FILING DATE: 1995-07-31
PRIOR APPLICATION NUMBER: 08/228,944
PRIOR FILING DATE: 1994-04-18
PRIOR APPLICATION NUMBER: 07/749,385
PRIOR FILING DATE: 1991-08-22
PRIOR APPLICATION NUMBER: 08/260,174
PRIOR FILING DATE: 1994-06-13
PRIOR APPLICATION NUMBER: 08/124,981
PRIOR FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 38
LENGTH: 716
TYPE: PRT
ORGANISM: human
US-10-040-430-38

Query Match 24.8% ; Score 114 ; DB 9 ; Length 716 ;
Best Local Similarity 34.7% ; Pred. No. 0.00011 ;
Matches 25 ; Conservative 19 ; Mismatches 26 ; Indels 2 ; Gaps 2 ;

QY 16 IRLKRRKSEDDAALIPGKLEICSDGCLSPRSRTSYVDGGLGSGYPLKDKSSGQPLSV 75
DB 331 VPKSRKTTLEQPPSVALKVEPVGEDLCSPPPADFAPEADYSSFOH-IRKGGFCDOYLAV 389


```

; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (248)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (273)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (323)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-940
```

```

Query Match          14.3% Score 65.5; DB 9; Length 557;
Best Local Similarity 28.8% Pred. No. 23;
Matches 23; Conservative 7; Mismatches 33; Indels 17; Gaps 3;
```

```

Qy 9 QVCVETDIPKTRKTSDDAAILPGKLEICSDDGNLSPSRETSDVDDGLGSGQ-----60
Db 446 QYVKEAKM---EKLKSNPATCTPGKPS--SKSQNKCPKPSQGLSTEENLSATITQPIHQ 500
```

```

Qy 61 ----YPLKKDSSGDOFLSVP 76
Db 501 KENIIPLVATNSDQFLTTP 520
```

```

RESULT 6
US-09-925-299-940
; Sequence 940, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 940
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (53)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (248)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (273)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (323)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

```

; LOCATION: (323)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-940
```

```

Query Match          14.3% Score 65.5; DB 10; Length 557;
Best Local Similarity 28.8% Pred. No. 23;
Matches 23; Conservative 7; Mismatches 33; Indels 17; Gaps 3;
```

```

Qy 9 QVCVETDIPKTRKTSDDAAILPGKLEICSDDGNLSPSRETSDVDDGLGSGQ-----60
```

```

Db 446 QYVKEAKM---EKLKSNPATCTPGKPS--SKSQNKCPKPSQGLSTEENLSATITQPIHQ 500
```

```

Qy 61 ----YPLKKDSSGDOFLSVP 76
Db 501 KENIIPLVATNSDQFLTTP 520
```

```

RESULT 7
US-09-405-920-2
; Sequence 2, Application US/09405920
; Publication No. US20030060435A1
; GENERAL INFORMATION:
; APPLICANT: CARILLO, Serge
; APPLICANT: BLANCHARD, Jean-Marie
; APPLICANT: PIECHACZYK, Marc
```

```

; TITLE OF INVENTION: METHOD OF CANCER TREATMENT BY P53 PROTEIN CONTROL
; FILE REFERENCE: ST94037A-US
```

```

; CURRENT APPLICATION NUMBER: US/09/405,920
```

```

; EARLIER FILING DATE: 1999-09-24
```

```

; EARLIER FILING DATE: 1996-11-27
```

```

; EARLIER APPLICATION NUMBER: FR94/06583
```

```

; EARLIER FILING DATE: 1994-05-31
```

```

; EARLIER APPLICATION NUMBER: WO PCT/FR95/00670
```

```

; EARLIER FILING DATE: 1995-05-22
```

```

; NUMBER OF SEQ ID NOS: 4
```

```

; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 2
```

```

; LENGTH: 694
```

```

; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
```

```

US-09-405-920-2
```

```

Query Match          14.1% Score 64.5; DB 9; Length 694;
```

```

Best Local Similarity 30.3% Pred. No. 39;
Matches 23; Conservative 9; Mismatches 25; Indels 19; Gaps 4;
```

```

Qy 13 ETDIPKTRKTSDDAAILPGKLEICSDDGNLSPSRETSDVDDGLGSGQYPLK--KDSGSD 70
```

```

Db 290 EPDLDRSIK-EVDEKAKKEKLEKCGED-----DETIPSEYRLKPATDCKGK 336
```

```

Qy 71 QFLSVSPPTWMSKPKP 86
```

```

Db 337 PLLPEPE---BKPKP 348
```

```

RESULT 8
US-09-764-864-1244
; Sequence 1244, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1244
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
```

FEATURE:
NAME/KEY: SITE
LOCATION: (16)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1244

Query Match 13.7%; Score 63; DB 10; Length 404;
Best Local Similarity 34.3%; Pred. No. 29;
Matches 24; Conservative 8; Mismatches 26; Indels 12; Gaps 3;

QY 13 ETDPLKRTKTSDDQAALPGLKLEICS---DDOGLSPSPRETSDVDGLSQYPLKDDSG 69
DB 21 ETQPPKRRRTIEDFN-----KFCSEYLAAYAGYIPSPKESDWPASGSSSPLRGESAA 73

QY 70 DQ--FLSVPS 77
DB 74 DSDGMDAPS 83

RESULT 9

US-09-765-272-120
Sequence 120, Application US/09765272
Patent No. US20020061545A1

GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 522 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 120:
US-09-765-272-120

Query Match 13.5%; Score 62; DB 10; Length 522;
Best Local Similarity 32.4%; Pred. No. 53;
Matches 22; Conservative 9; Mismatches 31; Indels 6; Gaps 3;

QY 9 QYCVETDPLKRTKTSDDQAALPGLKLEICSDQGLSPSPRETSDVDGLSQYPLKDD-- 66
DB 104 KYSADQDQPNSTKTSDD--VVHSADLE-WNQGQGVKSLQGEASGDGLSEKSSIAADNL 159

QY 67 SSGDQPLS 74
11111111

DB 160 SSNDSEFAS 167

RESULT 10

US-09-738-664-2
Sequence 2, Application US/09758664
Patent No. US20010025022A1

GENERAL INFORMATION:
APPLICANT: KIRLEY, KRISTINE KAY
APPLICANT: MICHAIOVICH, DAVID
TITLE OF INVENTION: HNOVILR
FILE REFERENCE: GP-70576-C1

CURRENT APPLICATION NUMBER: US/09/758,664
CURRENT FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: UK 97309517.7

PRIOR FILING DATE: 1997-11-26
PRIOR APPLICATION NUMBER: 09/187,711
PRIOR FILING DATE: 1998-11-06

NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2

LENGTH: 538
TYPE: PRT

ORGANISM: HOMO SAPIENS
US-09-758-664-2

Query Match 13.5%; Score 62; DB 10; Length 538;
Best Local Similarity 35.0%; Pred. No. 55;
Matches 35; Conservative 14; Mismatches 33; Indels 18; Gaps 9;

QY 1 LSPAPPPFOYCY--ETDPLKRTKTSDDQA--ALPGL--EICSDDGN---LSPSRE-T 51
DB 118 IKPAP-PENVVTYTFSGQYINISWRSDYEDPAFYMLGKLQYELQYHNRGPPMAYSPRRKLI 176

QY 52 SVDDGLGSOYPL--KKDSSGD--QPLSVSP-----FTMSK 83
DB 177 SYDSRSVSLHPLFRKDSYELQVRAGPMPGSSYQGTWSE 216

RESULT 11

US-09-765-272-118
Sequence 118, Application US/09765272
Patent No. US20020061545A1

GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

```

; INFORMATION FOR SEQ ID NO: 118:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 1040 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 118:
US-09-765-272-118

Query Match               13.5%; Score 62; DB 10; Length 1040;
Best Local Similarity     32.4%; Pred. No. 1.3e+02;
Matches 22: Conservative 9; Mismatches 31; Indels 6; Gaps 3;

QY 9 OYCETDIPKTRKTSDDOAILPGKLEICSDQGNLSPRRTSVDDGLGSGYPLKKD-- 66
DB 104 KYSARDSDGNPNSTKTS--VYHSADLE-WNCGCKVSLQGLASGDDGLSEKSIADNL 159
QY 67 SSGDOFLS 74
DB 160 SSNSDFAS 167

RESULT 12
US-09-965-529-20
; Sequence 20, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
;   APPLICANT: LAL, Precelt
;   APPLICANT: YUE, Henry
;   APPLICANT: TANG, Y. Tom
;   APPLICANT: BANDMAN, Olga
;   APPLICANT: BUREFORD, Neil
;   APPLICANT: AZIMZAI, Yalda
;   APPLICANT: BAUGHN, Mariah R.
;   APPLICANT: LU, Dyang Alma M.
;   APPLICANT: PATTERSON, Chandra
;   TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
;   FILE REFERENCE: PF-0731 USA
;   CURRENT APPLICATION NUMBER: US/09/965,529
;   PRIOR FILING DATE: 2001-09-26
;   PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
;   PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
;   NUMBER OF SEQ ID NOS: 74
;   SOFTWARE: PERL Program
;   SEQ ID NO 20
;   LENGTH: 427
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020182671A1 1656935CD1
US-09-965-529-20

Query Match               13.4%; Score 61.5; DB 9; Length 427;
Best Local Similarity     25.3%; Pred. No. 46;
Matches 20: Conservative 13; Mismatches 39; Indels 7; Gaps 2;

QY 13 EMDIPKTRKTSDDOAILPGKLEICSDQG-----NLSPRRTSVDDGLGSGYPLKKD 66
DB 21 GSRLPWMTIPNSHDNLPASASPLEONPSKHGALPGGLSTIGPGKSSIDSYG-RIDLION 79
QY 67 SSGDOFLSVSPFTWSKPK 85
DB 80 SESPASPVAVPHSWSRK 98

RESULT 13
US-10-159-151-2
; Sequence 2, Application US/10159151
; Publication No. US20030036075A1
; GENERAL INFORMATION:
;   APPLICANT: Cytokine169, Inc.
```

```

; APPLICANT: Beraud, Christopher
; APPLICANT: Guo, Jun
; APPLICANT: Freedman, Richard
; APPLICANT: Patel, Umesh A.
; APPLICANT: Davies, Katherine A.
;   TITLE OF INVENTION: KINI-3 MOTOR PROTEIN AND METHODS FOR ITS USE
;   FILE REFERENCE: 020552-002400US
;   CURRENT APPLICATION NUMBER: US/10/159,151
;   CURRENT FILING DATE: 2002-05-31
;   PRIOR APPLICATION NUMBER: US/09/967,908
;   PRIOR FILING DATE: 2001-09-26
;   PRIOR APPLICATION NUMBER: US 09/675,227
;   PRIOR FILING DATE: 2000-09-29
;   NUMBER OF SEQ ID NOS: 10
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 2
;   LENGTH: 1368
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Kln1-3 motor
;   OTHER INFORMATION: domain fragment
US-10-159-151-2

Query Match               13.4%; Score 61.5; DB 9; Length 1368;
Best Local Similarity     22.7%; Pred. No. 2e+02;
Matches 22: Conservative 14; Mismatches 22; Indels 39; Gaps 3;

QY 12 VETDIPKTRKTSDDOAILPGKLEICSDQGNLSPRRTSVDDGL----- 57
DB 893 VDSRDPINHRRAALDHS-----CSPSKGPVMSRSTSSGSPSPRSLAERPYCSQ 943
QY 58 -----GSGYPLKKDS-----SGDOFLSVSP 78
DB 944 VDFIYRQERNGGSSFDLRKDAQSEVSGENEGNLSP 980

RESULT 14
US-09-216-393-83
; Sequence 83, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
;   APPLICANT: Milhausen, Michael James
;   TITLE OF INVENTION: TOXOPLASMA GONDI PROTEINS, NUCLEIC ACID MOLECULES, AND
;   FILE REFERENCE: TX-1-C2
;   CURRENT APPLICATION NUMBER: US/09/216,393
;   CURRENT FILING DATE: 1998-12-18
;   EARLIER APPLICATION NUMBER: 08/994,825
;   EARLIER FILING DATE: 1997-12-19
;   NUMBER OF SEQ ID NOS: 364
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 83
;   LENGTH: 112
;   TYPE: PRT
;   ORGANISM: Toxoplasma gondii
US-09-216-393-83

Query Match               13.3%; Score 61; DB 10; Length 112;
Best Local Similarity     30.3%; Pred. No. 9.6;
Matches 27: Conservative 11; Mismatches 29; Indels 22; Gaps 5;

QY 1 LSPAPPPFOYCVETDIPKTRKTSDDOAILPGKLEICSDQGNLSPRRTSVDDGLGSG 60
DB 5 VSPSPF-----VSDSP-----SSPFAQSASPPP-----SEYDLSLPLAESY-----SS 45
QY 61 YPLKKDSSGDOFLSVSPFTWSK---PKP 86
DB 46 LPLAKQASPLHLTHQHPSPLMTQRASPS 74

RESULT 15
US-10-027-923-5
```

; Sequence 5, Application US/10027923
; Patent No. US2002014230A1
; GENERAL INFORMATION:
; APPLICANT: Briann Gaither Bates
; APPLICANT: Kamalaka Gulukota
; APPLICANT: Yunong Xie
; APPLICANT: Janet Elizabeth Paulsen
; TITLE OF INVENTION: NOVEL GLUTAMATE RECEPTOR MODULATORY PROTEINS AND
; FILE REFERENCE: GNN-024
; CURRENT APPLICATION NUMBER: US/10/027,923
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,589
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-027-923-5

Query Match 13.3%; Score 61; DB 12; Length 1203;
Best Local Similarity 29.9%; Pred. No. 2e+02; 38; Indels 6; Gaps 3;
Matches 23; Conservative 10; Mismatches 38; Indels 6; Gaps 3;
QY 5 PEPFOYCVETDIPDKTKRTSEDOAILPGKLEICSDDOG--NLSPSRETSDVDDGLGSOYP 62
Db 1088 PICSSYLIPKEIQPTMTTFEALIOPLPA-IEVTGAGAGATGVSPAQETPTG--AESAP 1143
QY 63 LKRDSSGDOFLSVSPF 79
Db 1144 GRPDLELVALTPSPF 1160

Search completed: April 11, 2003, 14:24:23
Job time : 17 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 18, 2003, 20:50:51 : Search time 2142 Seconds
(without alignments)
1168.460 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459

Sequence: 1 LSPAPFPFOYCVETDIPDKT.....SSCDPLSVSPFTWKKPKP 86

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlh
-O=/cgn2.1/USPRO.spool/US09550115/runat_11042003.141902.18148/app_query.fasta.1.263
-DB=GenEmbl -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09550115.ecgn.1.1.2496.grunat.11042003.141902.18148 -DEPU=6 -ICPU=3
-NO_XLPPX -NO_MMAR -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: gb_vl:*
17: gb_vl:*
18: gb_vl:*
19: gb_vl:*
20: gb_vl:*
21: gb_vl:*
22: gb_vl:*
23: gb_vl:*
24: gb_vl:*
25: gb_vl:*
26: gb_vl:*
27: gb_vl:*
28: gb_vl:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	459	100.0	3602	10 MMU28807	U28807 Mus muscu
2	459	100.0	3638	10 D85612	D85612 Mouse mRNA
3	450	98.0	80112	2 AC128800	AC128800 Rattus no
4	450	98.0	251055	2 AC094385	AC094385 Rattus no
5	430	93.7	2406	6 I38155	I38155 Sequence 7
6	430	93.7	2406	6 I80097	I80097 Sequence 7
7	430	93.7	2647	6 I38156	I38156 Sequence 9
8	430	93.7	2647	6 I80098	I80098 Sequence 9
9	430	93.7	3583	9 HSU85430	HSU85430 Homo sapien
10	430	93.7	3680	9 HSU14510	HSU14510 Homo sapien
11	430	93.7	3725	9 BC001050	BC001050 Homo sapien
12	430	93.7	3769	9 HSU85428	HSU85428 Homo sapien
13	430	93.7	3784	9 HSU85428	HSU85428 Homo sapien
14	430	93.7	3969	6 AX409566	AX409566 Sequence 11
15	430	93.7	3969	6 I38157	I38157 Sequence 11
16	430	93.7	3969	6 I80099	I80099 Sequence 11
17	430	93.7	3969	6 HUMFAT4A	HUMFAT4A Homo sapien
18	430	93.7	191565	9 HUMFAT4A	HUMFAT4A Homo sapien
19	430	93.7	201981	2 AC130462	AC130462 Homo sapi
20	430	93.7	201981	2 AC130462	AC130462 Homo sapi
21	199	42.3	4594	5 AB037471	AB037471 Xenopus l
22	129	28.1	167462	2 AC123532	AC123532 Mus muscu
23	129	28.1	167462	2 AC123532	AC123532 Mus muscu
24	126	27.5	3637	10 AF309388S1	AF309388 Mus muscu
25	126	27.5	4799	2 AC116083	AC116083 Rattus no
26	123	26.6	180194	2 AC116083	AC116083 Rattus no
27	122	26.6	2880	6 HUMFAT3A	HUMFAT3A Homo sapien
28	122	26.6	2881	6 I38154	I38154 Sequence 5
29	122	26.6	2881	6 I80096	I80096 Sequence 5
30	122	26.6	2937	6 BC008857	BC008857 Homo sapi
31	122	26.6	196533	9 CMS00YVI	CMS00YVI Human chr
32	114	24.8	2743	6 I38153	I38153 Sequence 3
33	114	24.8	2743	6 I80095	I80095 Sequence 3
34	114	24.8	2743	6 HSU08015	HSU08015 Human NF-AT
35	114	24.8	2749	6 AR058939	AR058939 Sequence
36	114	24.8	2749	6 AX469982	AX469982 Sequence
37	114	24.8	2751	6 AR105239	AR105239 Sequence
38	114	24.8	2751	6 AR119160	AR119160 Sequence
39	114	24.8	2751	6 AR123531	AR123531 Sequence
40	114	24.8	2751	6 AR138184	AR138184 Sequence
41	114	24.8	2751	6 AR176745	AR176745 Sequence
42	114	24.8	2935	6 HSU5736	HSU5736 Human trans
43	114	24.8	4617	9 HSU80918	HSU80918 Homo sapien
44	114	24.8	4678	9 HSU80917	HSU80917 Homo sapien
45	114	24.8	5502	6 AR198316	AR198316 Sequence

RESULT 1

ALIGNMENTS

DB:	10	Gaps:	0
US-09-550-115A-2 (1-86) x MMU28807 (1-3602)			
QY 1	LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20		
Db 931	CTCAGCCCTGCAGCGCTTCCATTCCTAGTACTGTGTAGAGACGACATCCCTTGAANA 990		
QY 21	ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysGluIleCysSerAsp 40		
Db 991	AGGAAGACTTCTGAAAGATCAAGCTGCTACTACCAAGAAATTAAGATCTGTTCAGAT 1050		
QY 41	AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60		
Db 1051	GATCAAGAGCAACTTATCCCTTCGCCGAGACATCAGTAGATGAGCGCTTGATCTCAG 1110		
QY 61	TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80		
Db 1111	TATCCTTTAAAGAAAGATCATCTGGTGCACCAATTTCTTCAGTCCCTTCACCCCTTTAC 1170		
QY 81	TyrSerLysProLysPro 86		
Db 1171	TGGAGCAACCAAGCCCT 1188		
RESULT 2			
LOCUS	D85612	3638 bp	linear ROD 06-FEB-1999
DEFINITION	Mouse mRNA for NFATx, complete cds.		
VERSION	D85612		
KEYWORDS	D85612.1 GI:1906311		
SOURCE	mnFATx1; NFATx; murine nuclear factor of activated T cell. Mus musculus thymic lymphoma cultured cell cell_line:EL-4 cDNA to mRNA.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (sites)		
TITLE	Liu, Y., Koyano-Nakagawa, N., Amasaki, Y., Saito-Obara, F., Ikeuchi, T., Imai, S., Takano, T., Arai, N., Yokota, T. and Arai, K. Calcineurin-dependent nuclear translocation of a murine transcription factor NFATx: molecular cloning and functional characterization		
JOURNAL	Mol. Biol. Cell 8 (1), 157-170 (1997)		
MEDLINE	97170074		
REFERENCE	2 (bases 1 to 3638)		
AUTHORS	Liu, Y., Koyano-Nakagawa, N., Amasaki, Y., Saito-Obara, F., Ikeuchi, T., Imai, S., Takano, T., Arai, N., Yokota, T. and Arai, K. Calcineurin Dependent Nuclear Translocation of a Murine Transcription Factor NFATx: Molecular Cloning and Functional Characterization		
TITLE	Unpublished		
JOURNAL	3 (bases 1 to 3638)		
REFERENCE	Liu, Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (28-MAY-1996) Jie Liu, Institute of Medical Science, The University of Tokyo, Department of Molecular and Developmental Biology, 4-6-1 Shirokanedai, Minato-Ku, Tokyo 108, Japan (E-mail: liujie@eng.ims.u-tokyo.ac.jp, Tel:81-3-5449-5664, Fax:81-3-5449-5424)		
JOURNAL	On Mar 25, 1997 this sequence version replaced gi:1842164.		
FEATURES	Location/Qualifiers		
SOURCE	1..3638		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/cell_line="EL-4"		
	/cell_type="cultured cell"		
	/tissue_type="thymic lymphoma"		
	1..3638		
	/gene="mnFATx1"		
	8..3235		
	/gene="mnFATx1"		
	/standard_name="murine nuclear factor of activated T cell"		
	/codon_start=-1		

```

/product="NFATx"
/protein_id="BA12833.1"
/db_xref="gi:1842165"
/translation="MTTANCAHDELDFKLVFGEDGAPAPPSPGSRPADLEPPDCAST
YFENVDPSTLTTPCLPHNGLPSHSSVLSFQLOSHKVEGECETLESYSPLG
KPECPESIQFTTSPNCOQLDAHEDDLQINDERELEPSPRHLPLEPSPRG
SLSPASSISSRWFSDASCSLSHYDDVDSELENAARFTGLSPGSPG
CPGSESHMOQYSGSLSLPSOSPCSPSSITDEMLSPRASPSSPTSGCKRH
SSAEVCYAGSLSPHSPSPSPSGSVTEDTWTLTAVHSGLSLSPAPPEOCYET
DPLKTRKTSDEDAIILPGLKELTSCDDOGNLSPEKRETSVDGLCSQYFLKDDSCDF
LSVSPFTWSKPKGHPHPIERTSLPRLDMLPPIFGCCLEKTEVQPTTHRAHETE
GSRKAVASTGHPVLLGLYSEKRNLMFIGNADRYLPHAFYQVHRTGTVAT
ASQETIISTVLEILPLNNMSASIDACILKRNIDILKRGEDIGKRNTRVL
VERVHIPOPSKVLISQIASIPVECSORSAOELPIREKYSINSCVNGHMTYGSN
FLPSKIFLEKGGDGPHEVEKILREKOGAHIVLEVPYHNPATSAVOYHFL
CNGRRKKSQSORFTTTPVLMKQEDREDTDLVSFSLPYPHSAQORPSESEGHPRDA
MSAPGLCQVOPATSKVASTHLCPOEDPVLFOODALSSLVNQCPLPDPSPSI
GSSYQISQTSWNGPTSLPVPASSQEDPVLFOODALSSLVNQCPLPDPSPSI
NSDATGHLAHSPIHVSQTPHLSQKVCYHCSNAGOTLSPADITGQPSHLPTTY
CPHSGSATASPASHPLASSPISGSPSLOPMPYQSPSSGATSPSSPTTMRHSGO
HSTQASTGCGGSLVPSVLSVCHSLCDPASFPFGATVSIKREPEDEGNFATIGLQDI
TLDDVNEITIGDMDSISVSQATEVWRDPLPSPASPLMTSHSH"
3638
/polyA_site
/gene="NFATx1"
/feature="18 A nucleotides"
BASE COUNT 960 a 944 c 737 g 997 t
ORIGIN
Alignment Scores:
Pred. No.: 1,42e-42 Length: 3638
Score: 459.00 Matches: 86
Percent: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-550-115a-2 (1-86) x DB5612 (1-3638)
Qy 1 LeuserProAlProPherProhegIntYrCyValGluThrAspIleProLeuLysThr 20
Db 968 CTCAGCCCTGCACCTTTTCATTTGATCTGTGAGACATGACATCCCTTTGAACA 1027
Qy 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
Db 1028 AGCAAGACTTCTGAGATCAACACTGCCATACACGAAATTAAGAGATCTGTTCAGAT 1087
Qy 41 AapGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlnSerGln 60
Db 1088 GATCAAGGAGAACTTATCCCTTCCCGGAGACATCATGATGATGCGCTTGATCTTCAG 1147
Qy 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1148 TATCCTTTAAAGAAATTCATCTGTGACCAATTTCTTCACTTCTTCACTTACC 1207
Qy 81 TrpSerLysProLysPro 86
Db 1208 TCGAGCAACAACCAAGCCT 1225
RESULT 3
LOCUS AC128800 80112 bp DNA linear HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-149113, *** SEQUENCING IN PROGRESS
*** 47 unordered pieces.
ACCESSION AC128800
VERSION AC128800.1 GI:21952659
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 80112)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

```

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimake,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathore,S.R., David,R.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Donthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eauenthal,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabels,A., Gao,D., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hughes,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lounsged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maneshwar,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkenko,S., Ogun,M., Okunonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabor,P., Tamerias,A., Tamerias,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleciyk,R., Wooten,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Glbbs,R.

Direct Submission
Unpublished
2 (bases 1 to 80112)
Worley,K.C.
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAMD
Center clone name: CH230-149113
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 45897 bases at least Q40
Consensus quality: 49666 bases at least Q30
Consensus quality: 52006 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1167: contig of 1167 bp in length
1168 1267: gap of unknown length
1268 2322: contig of 1055 bp in length
2323 2422: gap of unknown length
2423 3545: contig of 1123 bp in length
3546 3645: gap of unknown length
3646 4684: contig of 1039 bp in length
4685 4784: gap of unknown length
4785 5931: contig of 1147 bp in length
5932 6032: gap of unknown length
6032 7197: contig of 1166 bp in length
7198 7298: gap of unknown length
7298 9109: contig of 1812 bp in length
9110 10351: contig of 1142 bp in length
10352 11686: gap of unknown length
11687 11786: gap of unknown length
11787 13724: contig of 1938 bp in length
13725 13824: gap of unknown length
13825 15260: contig of 1436 bp in length
15261 15360: gap of unknown length
15361 16528: contig of 1168 bp in length
16529 16628: gap of unknown length
16628 17885: contig of 1267 bp in length
17885 17995: gap of unknown length
17995 19015: contig of 1020 bp in length
19016 19115: gap of unknown length
19116 20340: contig of 1225 bp in length
20341 22641: contig of 2201 bp in length
22642 22741: gap of unknown length
22741 24172: contig of 1431 bp in length
24172 24272: gap of unknown length
24273 25464: contig of 1192 bp in length
25465 25564: gap of unknown length
25565 26720: contig of 1156 bp in length
26721 26820: gap of unknown length
26821 27851: contig of 1031 bp in length
27852 27951: gap of unknown length
27952 29978: contig of 2027 bp in length
29979 30078: gap of unknown length
30079 31265: contig of 1187 bp in length
31266 31365: gap of unknown length
31366 32474: contig of 1109 bp in length
32475 32574: gap of unknown length
32575 34748: contig of 2174 bp in length
34749 34848: gap of unknown length
34849 35869: contig of 1021 bp in length
35870 35969: gap of unknown length
35970 36971: contig of 1002 bp in length
36972 37071: gap of unknown length
37072 39138: contig of 2067 bp in length
39139 39238: gap of unknown length
39239 40675: contig of 1437 bp in length
40676 40775: gap of unknown length
40776 41778: contig of 1003 bp in length
41779 41878: gap of unknown length
41879 43233: contig of 1355 bp in length
43234 43333: gap of unknown length
43334 44871: contig of 1538 bp in length
44872 44971: gap of unknown length
44972 46882: contig of 1911 bp in length
46883 48058: contig of 1076 bp in length
48059 48158: gap of unknown length
48159 49724: contig of 1566 bp in length
49725 49824: gap of unknown length
49825 52145: contig of 2321 bp in length
52146 52245: gap of unknown length
52246 54449: contig of 2204 bp in length
54450 54549: gap of unknown length
54550 55854: contig of 1305 bp in length

55855 55954: gap of unknown length
55955 57587: contig of 1633 bp in length
57588 57687: gap of unknown length
57688 60364: contig of 2677 bp in length
60365 60464: gap of unknown length
60465 62367: contig of 1903 bp in length
62368 62467: gap of unknown length
62468 64650: contig of 2183 bp in length
64651 64751: gap of unknown length
64751 66564: contig of 1814 bp in length
66565 66664: gap of unknown length
66665 69184: contig of 2520 bp in length
69185 69284: gap of unknown length
69285 71426: contig of 2142 bp in length
71427 71526: gap of unknown length
71527 75215: contig of 3689 bp in length
75216 75316: gap of unknown length
75316 77221: contig of 1906 bp in length
77222 77321: gap of unknown length
77322 80112: contig of 2791 bp in length.

FEATURES
source 1..80112 Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-149113"

BASE COUNT 18233 a 18609 c 21607 g 16798 t 4865 others
ORIGIN

Alignment Scores:
Pred. No.: 3 52e-40 Length: 80112
Score: 450.00 Matches: 83
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 96.51% Mismatches: 0
Query Match: 98.04% Indels: 0
DB: 2 Gaps: 0

US-09-550-115A-2 (1-86) x AC128800 (1-80112)

QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuYsrThr 20
Db 38499 CTCAGCCCTGCACCGCTTCATTCAGTACTGTAAGTGTGAGACATGACATCCCTTGAANAACA 38440

QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
Db 38439 AGGAAGACTTCACCATCAAGCTGCCATACACGAGAAATTTAGAGGCTGCTTCAGAT 38380

QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspGlyLeuGlySerGln 60
Db 38379 GATCAAGGAGAGCTTATCCCATCCCGGAGACATCAGTATGATGATGCGCTTGATCTCAG 38320

QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 38319 TATCCTTTAAGAAAGATTCATCTGCTGACCAATTTCTTCACTTCCCTTACCTTTACC 38260

QY 81 TTPSerLysProLysPro 86
Db 38259 TGGAGCAACAACAAAGCCT 38242

RESULT 4
LOCUS AC094385 251055 bp DNA linear HTG 10-JUL-2002
DEFINITION Rattus norvegicus clone CH230-3n11, *** SEQUENCING IN PROGRESS ***
ACCESSION AC094385
VERSION AC094385.4 GI:21716525
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Notway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 251055)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alshbrook, S. L., Amaralunga, H. C., Are, J. R., Ayele, M., Banks, T.,
Barber, J., Benton, J., Blamege, K., Blankenburg, K., Bonnin, D.,
Bouch, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N. P.,
Buahy, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,
Carroll, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,
Davila, M. L., Davis, C., Davy-Carroll, L., Dedertch, D. A.,
Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Dihn, H. J.,
Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, R. H.,
Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabbis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorell, J. H., Guevara, W., Gunnarone, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsoson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,
Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
Lo, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
Lozad, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maneshvari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, A., Mahoney, E., McLeod, M. P., Meador, M., Mel, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, N., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokkenko, S., Ogun, M., Okunolu, G.,
Otagunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Plims, E., Pu, L. L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtrati, N., Sisson, I.,
Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansay, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umanli, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

2 (pages 1 to 251055)
Direct Submission
Unpublished
2 (pages 1 to 251055)
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (pages 1 to 251055)
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT

Alshrooks, J.L., Amaraturung, H.C., Are, J.R., Ayele, A.M., Banks, T.,
Barbarel, J., Benton, J., Bimge, K., Blankenburg, K., Bonnin, D.,
Bouck, V., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathore, S.R., David, R.,
Davilla, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.V., Draper, H., Dugan-Kocha, S., Durbin, K.J.,
Earheart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorelli, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Horn, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Jarlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kraevic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louised, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Matinez, E.,
Massey, E., Maunula, E., McLeod, M.P., Medora, M., Mel, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, J., Monabbat, K., Morgan, M., Morris, S.,
Mosier, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nockemko, S., Ogih, M., Okunou, G.,
Oragunye, N., Olickson, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peterson, L., Pickens, R., Pitman, E., Pullis, M., Ren, Y.,
Rivers, M., Rojas, A., Rotundo, I., Rolfe, M., Ruiz, S., Savery, G.,
Scheer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, T.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Wainstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 251055)
Worley, K.C.

Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 251055)
Worley, K.C.

Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 9, 2002 this sequence version replaced gi:17941108.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GAMP
Center clone name: CH230-3N11
Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 204355 bases at least Q40
Consensus quality: 208911 bases at least Q30
Consensus quality: 212278 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)

* NOTE: This is a 'working draft' sequence. It currently	
* consists of 52 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
1	1108: contig of 1108 bp in length
1109	1108: gap of unknown length
1209	2683: contig of 1475 bp in length
2684	2783: gap of unknown length
2784	3883: contig of 1100 bp in length
3884	3883: gap of unknown length
3984	5519: contig of 1536 bp in length
5520	5519: gap of unknown length
5620	7019: contig of 1400 bp in length
7020	7119: gap of unknown length
7120	8194: contig of 1075 bp in length
8195	8294: gap of unknown length
8295	9646: contig of 1352 bp in length
9647	9746: gap of unknown length
9747	1088: contig of 1242 bp in length
10989	11088: gap of unknown length
11089	12378: contig of 1290 bp in length
12379	12478: gap of unknown length
12479	13746: contig of 1268 bp in length
13747	13846: gap of unknown length
13847	15479: contig of 1633 bp in length
15480	15579: gap of unknown length
15580	17146: contig of 1567 bp in length
17147	17746: gap of unknown length
17747	18599: contig of 1453 bp in length
18700	18799: gap of unknown length
18800	20609: contig of 1810 bp in length
20610	20709: gap of unknown length
20710	22107: contig of 1398 bp in length
22108	22207: gap of unknown length
22208	23709: contig of 1502 bp in length
23710	23809: gap of unknown length
23810	24912: contig of 1003 bp in length
24913	24912: gap of unknown length
24913	27254: contig of 2342 bp in length
27255	27254: gap of unknown length
27355	29236: contig of 1882 bp in length
29237	29336: gap of unknown length
29337	31709: contig of 2373 bp in length
31710	31809: gap of unknown length
31810	33833: contig of 2024 bp in length
33834	33933: gap of unknown length
33934	36015: contig of 2082 bp in length
36016	36115: gap of unknown length
36116	38441: contig of 2326 bp in length
38442	38541: gap of unknown length
38542	42667: contig of 4226 bp in length
42668	42867: gap of unknown length
42868	44532: contig of 1765 bp in length
44633	44732: gap of unknown length
44733	47340: contig of 2608 bp in length
47341	47440: gap of unknown length
47441	49374: contig of 1834 bp in length
49375	49374: gap of unknown length
49375	54375: contig of 5258 bp in length
54533	54732: gap of unknown length
54733	58371: contig of 3639 bp in length
58372	58471: gap of unknown length
58472	61803: contig of 3332 bp in length
61804	61903: gap of unknown length
61904	66184: contig of 4281 bp in length
66185	66284: gap of unknown length
66285	69862: contig of 3578 bp in length
69863	75962: gap of unknown length
69963	75962: contig of 5862 bp in length

* 75925 75924: gap of unknown length
* 75925 81868: contig of 5944 bp in length
* 81869 81968: gap of unknown length
* 81969 87828: contig of 5860 bp in length
* 87829 87928: gap of unknown length
* 94013 94112: gap of 6084 bp in length
* 99920 99920: contig of 5808 bp in length
* 100021 100020: gap of unknown length
* 105508 105508: contig of 5488 bp in length
* 105509 105608: gap of unknown length
* 105609 111533: contig of 5925 bp in length
* 111534 111633: gap of unknown length
* 111634 117749: contig of 6116 bp in length
* 117750 117849: gap of unknown length
* 117850 125485: contig of 7636 bp in length
* 125486 125585: gap of unknown length
* 125586 132663: contig of 7078 bp in length
* 132664 132763: gap of unknown length
* 132764 139566: contig of 6803 bp in length
* 139567 139666: gap of unknown length
* 149287 149287: contig of 9621 bp in length
* 149288 149387: gap of unknown length
* 159110 159110: contig of 9723 bp in length
* 159111 159210: gap of unknown length
* 159211 166473: contig of 7263 bp in length
* 166473 166573: gap of unknown length
* 177703 177703: contig of 11130 bp in length
* 177804 177803: gap of unknown length
* 188488 188488: contig of 10685 bp in length
* 188489 188588: gap of unknown length
* 199293 199293: contig of 10705 bp in length
* 199294 199393: gap of unknown length
* 211045 211045: contig of 11652 bp in length
* 211046 211145: gap of unknown length
* 230441 230441: contig of 19296 bp in length
* 230541 230541: gap of unknown length
* 230542 230542: contig of 20514 bp in length.

FEATURES
source
Location/Qualifiers
1..251055
/organism="Rattus norvegicus"

Alignment Scores:
Pred. No.: 1,13e-39 Length: 251055
Score: 450.00 Matches: 83
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 96.51% Mismatches: 0
Query Match: 98.04% Indels: 0
DB: 2 Gaps: 0

US-09-550-115a-2 (1-86) x AC094385 (1-251055)

QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
DB 46000 CTCAGCCCGCAGACCGTTTCATTTCAGTACTGTGTAGACAGTACATCCCTTTGAAAAA 46059
QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 46060 AGGAAGACTTCGACGATCAAGCTGCCATACCAACGAAATTAGAGCTGTCTCAGAT 46119
QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
DB 46120 GATCAAGGAGCTTATCCCATCCCGGAGACATCAGTGTAGATGAGCCCTTGATCTCAG 46179
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 46180 TATCCTTTAAAGAAATTCATCTGTGTGACCAATTTCTTCACTTCCTCACCCTTTTACC 46239
QY 81 TrpSerLysProLysPro 86
DB 46240 TGGAGCAAAACCAAGCCT 46257
RESULT 5

I38155
LOCUS I38155 2406 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 7 from patent US 5612455.
ACCESSION I38155
VERSION I38155.1 GI:2086145
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2406)
AUTHORS Hoey,T.
TITLE Nuclear factors and binding assay
JOURNAL Patent: US 5612455-A 7 18-MAR-1997;
FEATURES Location/Qualifiers
source
1..2406
/organism="unknown"

BASE COUNT 645 a 612 c 535 g 614 t
ORIGIN

Alignment Scores:
Pred. No.: 1.89e-39 Length: 2406
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
DB: 6 Gaps: 0

US-09-550-115a-2 (1-86) x I38155 (1-2406)

QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
DB 1171 CTGGCCCTGCGATTTTCATTTCAGTACTGTGTAGACAGTACATCCCTTCAAAA 1230
QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 1231 AGGAAGACTTCGACGATCAAGCTGCCATACCAACGAAATTAGAGCTGTCTCAGAT 1290
QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
DB 1291 GACCAAGGAGGATTATCACACCCCGGAGACTTCATATGATGATGAGGCTTGATCTCAG 1350
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 1351 TATCCTTTAAAGAAATTCATCTGTGTGACCAATTTCTTCACTTCCTCACCCTTTTACC 1410
QY 81 TrpSerLysProLysPro 86
DB 1411 TGGAGCAAAACCAAGCCT 1428
RESULT 6

180097 2406 bp DNA linear PAT 10-JUN-1998
LOCUS 180097
DEFINITION Sequence 7 from patent US 5708158.
ACCESSION 180097
VERSION 180097.1 GI:3208387
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2406)
AUTHORS Hoey,T.
TITLE Nuclear factors and binding assays
JOURNAL Patent: US 5708158-A 7 13-JAN-1998;
FEATURES Location/Qualifiers
source
1..2406
/organism="unknown"

BASE COUNT 645 a 612 c 535 g 614 t
ORIGIN

Alignment Scores:
Pred. No.: 1.89e-39 Length: 2406
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4

FEATURES	source		
JOURNAL MEDLINE PUBLISHED	Immamura, R., Masuda, E.S., Naito, Y., Imai, S., Fujino, T., Takano, T., Arai, K. and Arai, N.		
REFERENCE	Cardoxyl-terminal 15-amino acid sequence of NFATx1 is possibly created by tissue-specific splicing and is essential for transactivation activity in T cells		
AUTHORS	J. Immunol. 161 (7), 3455-3463 (1998)		
TITLE	98430664		
REFERENCE	9759864		
TITLE	3 (bases 1 to 3583)		
REFERENCE	Arai, N.		
TITLE	Direct Submission		
REFERENCE	Submitted (15-JUN-1997) Cell Signaling, DNAX Research Institute, 901 California Ave, Palo Alto, CA 94304-1104, USA		
FEATURES	Location/Qualifiers		
source	1. 3583		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/cell_line="T antigen-transformed diploid fibroblast"		
	/cell_type="fibroblast"		
	25. 3162		
	/note="Isoform of NFATx"		
	/codon_start=1		
	/product="transcription factor NFATx4"		
	/protein_id="AAB46597.1"		
	/db_xref="gi:1835591"		
	/translation="MTTANCGAHDELDFKLVFGEDGAPAPPFGSGRPADLEDDPCASTIYIENVDPSPSTLTPRLCLPHHGILPSHSVSYSPSQLOSHKNEBETCIPESKIYPLDGGKPEPCPSIQITTSIPMKOHELDHEDLDQINDERELEPSPKHDLIPLLESTRESKLSPPASSISSRSWFSFDASSCESLHIYDDVDELENAARFLGSLPLSPGSPGSGPGSGEETWAOYGLGHSLSPPQSPCHSPRSSVTDENMLSPRASGSPSSRPSPCGKRHSASAEYCHGSLSPHSHPVSPSGHPGSGVTEDTWLNASVHGSGSLGAPVFPQYCVETEDLPKTRKTSIEDQALIPGKLEICSDOGSLSPARETISIDGLSGLOPKKSDSGDOELVSPPTMSKPKRGKHPHPTSPSLDPLMLPAHFGOCELIKIEVOKTHHRAHYEFEGSRVAVASTGCHPVYKLGKNEPNIQMTGADRYLRPAHYVNHRTGTVATASQETIIASTVLEIPLPERNMASITDCAQELILNNSDELKGEYDIGKTRVTLVERHIIQPSGKVLIDIASIPVECSQSAQDELPIEKYSINSCSVNGHEMVTGSENFLPFSKIIIFLEKGGDGRPOMEVEKIIIREKQGNHIVLEVPYINPVAIVAOVHEFLCNGKRKKSQSORFETFTYVLLKQEHREEDILDSVSLVPHPAQORPSDGSCHDSVLSGQRLSIPQYASMTWTSNHLPOLCCRESKYSQKHMLPSPIVHQPVPPTPVYGGSPQPMOTNVYVNGPCIPINASSSEOPSSVMEQDQATLGVNLGCOLPASTPHFSNNGSGTHLAPHPSVHTLPHIOMSKYHCSNCGQRLSSVSGQITIGQSSLOLPTTYGPHSGSSVTLAPSAASHPLGSSSPFOFPMFYQSSSGTSSGSPATMHSGGQHSYQASQSGGGLSAPSSLICHSLEDCPASPDPGATVSIKPEDEDEPNPATIGDDITLDDGKFIISMFJX"		
BASE COUNT	959 a 898 c 757 g 969 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.83e-39		
Score:	430.00		
Best Similarity:	96.51%		
Best Local Similarity:	91.86%		
Query Match:	93.68%		
Db:	9		
	Gaps: 0		
US-09-550-115A-2 (1-86) x HSU85430 (1-3583)			
Qy	1	leuserproalaprohphepophedlnrycyysvalaigutmraspilleeproleulysrthr	20
Db	985	CTTGCCCTCGAGAGTTTTCATTCATTCAGACGTGTAGAGACTGACATCCCTCTCAAAAACA	1044
Qy	21	ArglystrSerGluAspGlnAlaIalaleuuproglyLysLeuGlnIleCysSerAsp	40
Db	1045	AGGAAACTTCTGGAAGATCAACGTCCTACTACCAGSAAAAATTAGAGCTGTGTTACAGT	1104
Qy	41	AspGlnGlyAsnLeuSerProSerArgLutHrSerValaAspAspGlyLeuGlySerGln	60
Db	1105	GACCAAGGAGATTATACACGACCGGAGACTCAATATAGATGATGCGCTTGATCTCAG	1164
Qy	61	TyrProLeuLysLysAspSerSerGlyLysGlnPheLeuSerValProSerProPheThr	80
Db	1165	TATCCTTTAAAGAAAGATTATGTCGTATCATAGTTTCTTTCAGTTCACCTTCACCTTTAC	1224

QY	81	7TPseTfvsProLyapPro	86	
Db	1225	TGGAGCAACCAAGCCT	1242	
RESULT 10				
LOCUS	HSU14510		3680 bp	mRNA
DEFINITION	Human transcription factor NFATx mRNA, complete cds.			PRI 31-JAN-1996
ACCESSION	U14510			
VERSION	U14510.1	GI:780373		
KEYWORDS				
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 3680) Masuda,E.S., Naito,Y., Tokumitsu,H., Campbell,D., Saito,F., Hannum,C., Arai,K. and Arai,N.			
TITLE	NFATx, a novel member of the nuclear factor of activated T cells family that is expressed predominantly in the thymus			
JOURNAL	Mol. Cell. Biol. 15 (5), 2697-2706 (1995)			
MEDLINE	95257951			
POBMED	77393550			
REFERENCE	2 (bases 1 to 3680)			
AUTHORS	Arai,N.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-SEP-1994) Naoko Arai, Molecular Biology, DNAX Research Institute of Molecular and Cellular Biology, 901 California Avenue, Palo Alto, CA 94304-1104, USA			
FEATURES	location/Qualifiers			
source	1..3680			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="16"			
	/map="16q21-q22"			
	/cell_line="Jurkat"			
	/cell_type="T-cell"			
	/note="in LambdaZipox vector"			
	25..3252			
	/function="transcription factor"			
	/codon_start=1			
	/product="NFATx"			
	/protein_id="AAA86308.1"			
	/db_xref="GI:780374"			
	/translation="MTTACGAGHDELDFKLVFGDGADAPPPGSPADLEPDDCAISI YIENVDPSPRTPTPCLPHHGGLPSHSVTSPPOLQSHKNTGTEIPEKSKSPGG KPEPCESIOTTISPNCHDELDHEDLDINDPERFRLTHSDRHLYLPDESYES SLSASPASSISRMSWESDASSCESLSHIYDDVSELENAARTFLHSPLTSPGSPFG CPGEGETHQOGLGHSISPRQSPCHSPRSSVTDENMLSPRPAGSPSRPSPGKRRH SSAVCYAGSLSPHHSPVSPGSHSPGSGSEDTMLNASVGGGGLGPAVFPQYCVET DLPKTRKTSDDOAILPGKLELCSDDOGLSPARETSIDDGAGSOYPLKDDGDOE LSVSPETWTKSPKPGHPTPIRTSTSLPRLDPAHNGCCLEKLEQPKTHRAHYETE GSRVAVASTGCHPVPVKKLLGNEKPTLQWFTGTDRIIRPAFTQVARTITKTVAT ASQGLIILASTKVLPLLPENNMNSASIDCAGLILRNSDLELRKGEDIGKRTVRAL VFRHINQPGSKVLESLQIASIPVCSORSKQELPHIEKYSINCSVNGGHEWVTGSN FLPEVKIIFLEGGDGRPOMEVEGKIIRKCCOAHILEVPRYHPNPAVAVOHFYL CNCKRRKSQSORETRYPLVLMKOEKREIDILSVSLDVPBAQTORPSSSGCSHVS LSGRSLICSIPTQYASMYTSHLPOLQCDDESYSKQNHITPPIYHQPOVYPTPPV GSSTQPMQNVNIVNGPTCLPINAASSQEDFVLFQODATLGLVNLGCPPLSIPTHS SNGSSTGLTAAHPVSHVTLPHLQSMGYHCSNKGORLSSPVADQITGGPSSQLOITP YGPHSSGATTAASHPASSPLSGSPOLQPMPOYQSPSGSTASSPAPRIMSG OHSTQAOSTGGGLSAPSLICSHLCSIDPASPAPPDGATVSIKPEPDEREPNATIGLD ITLDNVETIIGRMSQLSVSGAGYSGNAPLPSPSLIDGRSNGL"			
CDS				
BASE COUNT	948 a	926 c	781 g	1025 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2	91e-39	Length:	3680
Score:	430.00		Matches:	79
Percent Similarity:	96.51%		Conservative:	4
Best Local Similarity:	91.86%		Mismatches:	3
Query Match:	93.68%		Indels:	0

Alignment Scores:	
Pred. No.:	2,91e-39
Score:	430.00
Percent Similarity:	96.51%
Best Local Similarity:	91.86%
Query Match:	93.68%
Length:	3680
Matches:	79
Conservative:	4
Mismatches:	3
Indels:	0

US-09-550-115a-2 (1-86) x HSU14510 (1-3680)

Oy 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
 Db 985 CTTGGCCCTGACGTTTTCATTTTCAGTGTGTAGACTGACATCCCTCTCAAAACA 1044
 Oy 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
 Db 1045 AGGAAACTTGTGAAGATCAAGCTGCTACTACAGGAAATTAAGACCTGTTCAGAT 1104
 Oy 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
 Db 1105 GACCAAGGAGACTTATACACGACCGGAGACTCAATAGATGATGCCCTTGATCTCAG 1164
 Oy 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
 Db 1165 TATCCTTTAAAGAAAGATCATGTGTGATCATGTTCTTTCAGTTCCTTCACCTTAC 1224
 Oy 81 TrpSerLysProLysPro 86
 Db 1225 TGGAGCAACCAAGCCT 1242

RESULT 11
BC001050
LOCUS

DEFINITION Homo sapiens, nuclear factor of activated T-cells, cytoplasmic,
 calcineurin-dependent 3, clone MGC:1495 IMAGE:3505967, mRNA,
 complete cds.

ACCESSION BC001050
 VERSION BC001050.1 GI:12654448
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.

Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Primates: Catarrhini: Homiinae: Homo.
 1 (bases 1 to 3725)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (11-DEC-2000) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
 Greene, Mark Ketteman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 7 Row: d Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4758803.

FEATURES

source
 1. 3725
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /db_xref="taxon:9606"
 /clone="MGC:1495 IMAGE:3505967"
 /tissue_type="Kidney, renal cell adenocarcinoma"
 /clone_id="NIH_MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: POTB7"
 44. 3271
 /codon_start=1
 /product="nuclear factor of activated T-cells,

cytoplasmic, calcineurin-dependent 3"
 /protein_id="AA01050.1"
 /db_xref="GI:12654448"
 /translation="MTTANGAGHDELDFKLVGEDGAPAPPSPROSPRADLEPPDCASI
 YFENVDPPTLTPLCLPHHGLPSHSSVSPQLOSHNNYECTCIPESKYSPLGG
 KPRECPISQITISIPNQHQLDHAHEDLOINPERFELPNDHLYPLLEPSYRS
 SLSPSPASSISRSWFSADASSCELSHIYDQVDSLENAARFLGSLPSPGSPG
 CPGEETHQOQGLGHSLSRPSRCHSRSSVTDENMLSPRASPSSRRPSPGKRH
 SSAEVCYAGSLSPHSPSPGSPRCSSTEDMLNLSVHGSGSLSPAPFVOCVET
 DPLKPKRTSEDDAAILPKLCLSDSDQGLSPARETSTDDGLSOPPLKDKSGDQF
 LSVSPPTWSPKPKRGHPITRTSLPLDPLPPLHFGQCELEVOKTHHRAVETE
 GSRGAVNASTGHPVNLGYNENKPIQMFITGADRYLRPAFYOVHRTGCTVAVT
 ASOEIILASTKYLEILPLPENNMASISDCGILKRNDELKGETDIRKNTVRL
 VFRVHILPQSGKVLQIASIPVSCSRSAOELPIEKYSINSGVGHMAYGTGSN
 FLPESKTIFLEKODGRPOMEVEKTRREKCCGATLYLEVPRYNPAVTAAYVHPTL
 CNGKRKRSOSORFTYTPVLKOEIRETIDSSVSPSLVPPPAOTQRRSSDGSNDY
 LSGORSLGISTPQTASMTVSSHLPQLOCDSESKQHMIPSTVHPORVPTVPT
 GSSYQPMQTVNVVNGPTCLPINAASQSEFSPVLEQODATLGLVNLGCPILSPHS
 SNGSTGHLAHPHSAVHTLPHLOSMGYSNTOGRLSPVADQITGQPSLOPRTT
 YGPHSGSATTPASPAASHPLASSPLSGSPPOLOPMPYQSSSGTASSPAPTRMHS
 QHSTQAOASTGGGLSAPSSLCHSLICDPASEFPDPGATVSIKPEDEDEPNFATIGLD
 ITLDVNEIIGRMSQISVSGAGVSRQALPLPESLDRSGL"

BASE COUNT 979 a 929 c 791 g 1026 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,95e-39 Length: 3725
 Score: 430.00 Matches: 79
 Percent Similarity: 96.51% Conserved: 4
 Best Local Similarity: 91.86% Mismatches: 3
 Query Match: 93.68% Indels: 0
 DB: Gaps: 0

US-09-550-115a-2 (1-86) x BC001050 (1-3725)

Oy 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
 Db 1004 CTTGGCCCTGACGTTTTCATTTTCAGTGTGTAGACTGACATCCCTCTCAAAACA 1063
 Oy 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
 Db 1064 AGGAAACTTGTGAAGATCAAGCTGCTACTACAGGAAATTAAGACCTGTTCAGAT 1123
 Oy 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
 Db 1124 GACCAAGGAGACTTATACACGACCGGAGACTCAATAGATGATGCCCTTGATCTCAG 1183
 Oy 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
 Db 1184 TATCCTTTAAAGAAAGATCATGTGTGATCATGTTCTTTCAGTTCCTTCACCTTAC 1243
 Oy 81 TrpSerLysProLysPro 86
 Db 1244 TGGAGCAACCAAGCCT 1261

RESULT 12
 LOCUS HS085429 3769 bp mRNA linear PRI 03-MAY-2001
 DEFINITION Homo sapiens transcription factor NFATx3 mRNA, complete cds.
 ACCESSION U85429
 VERSION U85429.1 GI:1835588
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.

Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Primates: Catarrhini: Homiinae: Homo.
 1 (bases 1 to 3769)
 REFERENCE Masuda, E.S., Naito, Y., Tokumitsu, H., Campbell, D., Saito, F.,
 Hattori, C., Arai, K. and Arai, N.
 NFATx3, a novel member of the nuclear factor of activated T cells
 family that is expressed predominantly in the thymus
 JOURNAL Mol. Cell. Biol. 15 (5), 2697-2706 (1995)
 MEDLINE 95257951

```

PUBMED 7739550
REFERENCE 2 (bases 1 to 3769)
AUTHORS Imanura,R., Masuda,E.S., Naito,Y., Imai,S., Fujino,T., Takano,T.,
TITLE Carboxyl-terminal 15-amino acid sequence of NFATx1 is possibly
created by tissue-specific splicing and is essential for
transactivation activity in T cells
JOURNAL J. Immunol. 161 (7), 3455-3463 (1998)
MEDLINE 98430664
PUBMED 9759864
REFERENCE 3 (bases 1 to 3769)
AUTHORS Arai,N.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-1997) Cell Signalling, DNAX Research Institute,
901 California Ave, Palo Alto, CA 94304-1104, USA
FEATURES
source
location/Qualifiers
1..3769
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Jurkat"
/cell_type="T-cell"
25..3222
/note="isoform of NFATx"
/codon_start=1
/product="transcription factor NFATx3"
/protein_id="AAB46596.1"
/db_xref="GI:1835589"
/translation="MTTANCGAHDELDFKIVGEGDAPAPPPGSRPADLEPPDCASI
YIFNDDPSTLTPPLCLPHGLGHSVLSPPSLOSHKNEGCELPESKYSPLGG
KPRPCPSIOTITISIPNCHOEDLNDPREFLEERSRHLLTLPESYRES
SLSPSPASISRSWFSDSACESLSHITDYDSDLENAARFTLGSLTSPGSGPG
CPGETWQQVGLGHSLSRQSPCHSPSSVDEWMLPRPASGSSRPTSCGRRH
SSAEVYAGSLSPHNSPVPSPGSPGSDTEBTWLNASVHGSGIGPVPFOYCVET
DIPLEKRTSEDOAAILPGKLELCSDDGSLSPAETSIDLGGQYPLKDCSGDOR
LSVPEPTWPKRGKPHPTPIFRSSLPIDWMLPAHGGCELEKIEQPTNHHAYETE
GSRGVKASTGKHVPVYKLGLENERINLOMTGTADRYLPHAYOYHRTGKTVAT
ASOEIITASTKYLEIPLPENNMASIDCAILKRNSDIELRKGETDGRKNTVRL
VFRVHIPOPSGVLSLQIASIPVECSQSADELPIETKYSINSCVNGHEWVYGSN
FLPESKIIIFLEKGGDGRQWEVECKIIEKCGAHIVLEPPYHNPATVAIVHYEL
NGKRRKQSOSOFRTYTPVLLKOEHEEIDLSVSLPVPHAQOTQRPSSDCSHDSV
LSGSRKLSISIPOTYASWTSLSHLPOLCCRDSESVKEDHMLPSPIVHOPOVTPPPV
GSSYOPMOTNVVYNGPCLPTINAASSOEPDSVWEPQDNLGSLVNLGCPRLSTPFS
SNSGSGVTLTAHPHVSHTLPHLOSMGTHCSVTGORSLSPEVGDITGPPSLOLPT
YGPSSHSGVTTASPAASHPLGSSPLSGPPSPQFQPMYQSPSSGSSPSPATRMHSG
OHSIOASTGOGGLSPASSLSLCHSLCDPASPPDQATVSIKPEPDRPNFATIGLQD
ITLDDQCTSLDLEHQPSSAEKPMHVSILSCPAPFWRI"
BASE COUNT 974 a 939 c 802 g 1054 t
ORIGIN
Alignment Scores:
Pred. No.: 2.98e-39 Length: 3769
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB:
US-09-550-115a-2 (1-86) x HS085429 (1-3769)
QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
DB 985 CTTGGCCCTGCGAGTTTTCATTTCAGTACTGTGAGACGACGACATCCCTCCCAAAA 1044
QY 21 ArgLysThrSerGlnuAspGlnAlaIleLeuProGlyLysLeuGluLecysSerAsp 40
DB 1045 AGGAAAACCTCTGAGATACAGCTGCATACACGAGAAATAGAGCTGTGTTAGAT 1104
QY 41 AspGlnGlyAsnLeuSerProSerArgLutThrSerValAspAspGlyLeuGlySerGln 60
DB 1105 GACCAAGGAGATTATCACACGCGGAGACCTTCAATAGATGATGCGCTTGATCTCAG 1164
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80

```

```

DB 1165 TATCCTTTAAAGAAAGATTCATGCTGATCAGTTCTTTCAGTTCCCTTACC 1224
QY 81 TrpSerLysProLysPro 86
DB 1225 TGGAGCAACCAAGCCT 1242
RESULT 13
HS085428 3764 bp mRNA linear PRI 03-MAY-2001
LOCUS HS085428
DEFINITION Homo sapiens transcription factor NFATx2 mRNA, complete cds.
ACCESSION U85428
VERSION U85428.1 GI:1835586
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3784)
Masuda,E.S., Naito,Y., Tokumitsu,H., Campbell,D., Salto,F.,
Hannum,C., Arai,K. and Arai,N.
NFATx, a novel member of the nuclear factor of activated T cells
family that is expressed predominantly in the thymus
Mol.Cell. Biol. 15 (5), 2697-2706 (1995)
95257951
JOURNAL 7739550
MEDLINE 95257951
PUBMED
REFERENCE 2 (bases 1 to 3784)
AUTHORS Imanura,R., Masuda,E.S., Naito,Y., Imai,S., Fujino,T., Takano,T.,
Arai,K. and Arai,N.
TITLE Carboxyl-terminal 15-amino acid sequence of NFATx1 is possibly
created by tissue-specific splicing and is essential for
transactivation activity in T cells
J. Immunol. 161 (7), 3455-3463 (1998)
98430664
PUBMED 9759864
REFERENCE 3 (bases 1 to 3784)
AUTHORS Arai,N.
TITLE Direct Submission
JOURNAL Submitted (15-JAN-1997) Cell Signalling, DNAX Research Institute,
901 California Ave, Palo Alto, CA 94304-1104, USA
FEATURES
source
location/Qualifiers
1..3784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Jurkat"
/cell_type="T-cell"
25..3231
/note="isoform of NFATx"
/codon_start=1
/product="transcription factor NFATx2"
/protein_id="AAB46595.1"
/db_xref="GI:1835587"
/translation="MTTANCGAHDELDFKIVGEGDAPAPPPGSRPADLEPPDCASI
YIFNDDPSTLTPPLCLPHGLGHSVLSPPSLOSHKNEGCELPESKYSPLGG
KPRPCPSIOTITISIPNCHOEDLNDPREFLEERSRHLLTLPESYRES
SLSPSPASISRSWFSDSACESLSHITDYDSDLENAARFTLGSLTSPGSGPG
CPGETWQQVGLGHSLSRQSPCHSPSSVDEWMLPRPASGSSRPTSCGRRH
SSAEVYAGSLSPHNSPVPSPGSPGSDTEBTWLNASVHGSGIGPVPFOYCVET
DIPLEKRTSEDOAAILPGKLELCSDDGSLSPAETSIDLGGQYPLKDCSGDOR
LSVPEPTWPKRGKPHPTPIFRSSLPIDWMLPAHGGCELEKIEQPTNHHAYETE
GSRGVKASTGKHVPVYKLGLENERINLOMTGTADRYLPHAYOYHRTGKTVAT
ASOEIITASTKYLEIPLPENNMASIDCAILKRNSDIELRKGETDGRKNTVRL
VFRVHIPOPSGVLSLQIASIPVECSQSADELPIETKYSINSCVNGHEWVYGSN
FLPESKIIIFLEKGGDGRQWEVECKIIEKCGAHIVLEPPYHNPATVAIVHYEL
NGKRRKQSOSOFRTYTPVLLKOEHEEIDLSVSLPVPHAQOTQRPSSDCSHDSV
LSGSRKLSISIPOTYASWTSLSHLPOLCCRDSESVKEDHMLPSPIVHOPOVTPPPV
GSSYOPMOTNVVYNGPCLPTINAASSOEPDSVWEPQDNLGSLVNLGCPRLSTPFS
SNSGSGVTLTAHPHVSHTLPHLOSMGTHCSVTGORSLSPEVGDITGPPSLOLPT
YGPSSHSGVTTASPAASHPLGSSPLSGPPSPQFQPMYQSPSSGSSPSPATRMHSG
OHSIOASTGOGGLSPASSLSLCHSLCDPASPPDQATVSIKPEPDRPNFATIGLQD
ITLDDQCTSLDLEHQPSSAEKPMHVSILSCPAPFWRI"
BASE COUNT 979 a 950 c 804 g 1051 t
ORIGIN

```


GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: April 18, 2003, 20:52:51 ; Search time 235 Seconds

(without alignments)
824.136 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459
Sequence: 1 LSPAFPPQXCVETDIPLKT.....SSGDFLSVPSPFTWSPKPKP 86

Scoring table:

BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=x11
-Q/cgcn2.1/USPTO.spool/US09550115-runatc.11042003.141902.18138/app.query.fasta.1.263
-DB=N.Geneseq.101002 -OPMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09550115_cgcn.1.1.263-runatc.11042003.141902.18138 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONCLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database : N.Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	430	93.7	2406	17	AAT33677	Human transcriptio
2	430	93.7	3583	21	AAF21183	Human low adenosin
3	430	93.7	3583	21	AAA35061	Human adenosine re
4	430	93.7	3769	21	AAF21182	Human low adenosin
5	430	93.7	3769	21	AAA35060	Human adenosine re
6	430	93.7	3784	21	AAF21181	Human low adenosin
7	430	93.7	3784	21	AAA35059	Human adenosine re
8	430	93.7	3969	24	ABN95715	Human adenosine re
9	430	93.7	4010	16	AAO84303	Gene #2213 used to
10	430	93.7	17069	21	AAF21185	Human NF-AT120, X
11	430	93.7	17069	21	AAA35063	Human low adenosin
12	187	40.7	240	21	AAC14360	Human adenosine re
13	122	26.6	2881	17	AAT36868	Human secreted pro
14	122	26.6	2881	20	AAOX08717	Human transcriptio
15	114	24.8	2743	17	AAT36867	NF-AT3 encoding se
16	114	24.8	2743	22	AAH02897	Human transcriptio
17	114	24.8	2749	16	AAO86687	Human shear stress
18	114	24.8	2750	21	AAO86687	Human NF-AT transc
19	114	24.8	2751	22	AAAF31710	Human nuclear fact
20	114	24.8	2751	24	AAD34048	Human NF-ATC CDNA.
21	114	24.8	2853	16	AAO84301	Human cytoplasmic
22	114	24.8	2935	21	AAF21184	Human low adenosin
23	114	24.8	2935	21	AAA35062	Human adenosine re
24	114	24.8	4963	23	AA578396	DNA encoding novel
25	114	24.8	5502	24	AAD34036	Human NF-ATC CDNA.
26	79.5	17.3	4370	24	ABK84613	Human cytoplasmic
27	77	16.8	4803	23	ABL20913	Human cDNA differe
28	77	16.8	4803	23	ABL16961	Drosophila melanog
29	77	16.8	6225	23	ABL20912	Drosophila melanog
30	77	16.8	7027	23	ABL15960	Drosophila melanog
31	76	16.6	32768	20	AAK12954	Human prostatic exp
32	75.5	16.4	3825	23	AAV25634	S. cellulosum DNA
33	73	15.9	37856	21	AAV11992	DNA encoding novel
34	71.5	15.6	1257	23	AA5808302	DNA encoding novel
35	71.5	15.6	1257	23	AA580830	DNA encoding novel
36	71.5	15.6	1257	23	AA581402	DNA encoding novel
37	71.5	15.6	2223	22	AA500623	Human death-associ
38	71.5	15.6	2271	23	AA578101	Human Fas-binding
39	71.5	15.6	2477	19	AAV45136	Human Fas-binding
40	71.5	15.6	2477	22	AAAF72902	Human Fas-binding
41	71.5	15.6	2682	21	AAA88863	Human dentin matre
42	71.5	15.6	2811	23	AA580952	DNA encoding novel
43	71	15.5	3011	23	ABL07090	Drosophila melanog
44	71	15.5	3011	23	ABL07130	Drosophila melanog
45	71	15.5	32189	22	AA530115	Human lung antigen

ALIGNMENTS

RESULT 1
AAT33677
ID AAT33677 standard; CDNA: 2406 BP.

AC AAT33677:

DT 18-NOV-1996 (first entry)

DE Human transcription factor NFAT4a CDNA.

KW Nuclear factor of activated T-cells; NFAT; NFAT4a;

KW transcription factor; binding assay; immune system disease;

OS gene therapy; diagnosis; ds.

XX Homo sapiens.

XX FT

XX CDS

Location/Qualifiers
211..2337

```
FT      /tag- a
FT      /product- NFAT4a
XX
XX      MO9626959-A1.
PN
XX
XX      06-SEP-1996.
PD
XX
XX      04-MAR-1996; 96MO-US03113.
PF
XX
XX      02-MAR-1995; 95US-0396479.
PR
XX
XX      (TULA-) TULARIK INC.
PA
XX
XX      Hoey T;
PI
XX
XX      WPI: 1996-412738/41.
DR
XX      P-PSDB; AAM02251.
PT
XX      DNA mol. encoding human nuclear factors of activated T cells -
PT      useful for screening potential therapeutic and diagnostic agents for
PT      immune system diseases
XX
XX      PS      Disclosure: Page 50-52; 64pp; English.
XX
XX      A cDNA clone (AAT33677) codes for human nuclear factor of activated
XX      T-cells class 4a, NFAT4a (AAM02251), a regulator of immune system
XX      function. 2 Other classes of NFAT4 cDNAs, NFAT4b and NFAT4c,
XX      CC      were also isolated (see also AAT36975 and AAT36871) that differed from
XX      CC      NFAT4a at their 3' terminal sequences: the NFAT4c transcript was
XX      CC      most abundant. The major sites for NFAT4 expression are in skeletal
XX      CC      muscle and thymus. NFAT cDNAs (see also AAT36866-68) were isolated
XX      CC      from cDNA libraries prep. from Jurkat T-cells and human peripheral
XX      CC      blood lymphocytes by screening with a probe derived from the rel
XX      CC      similarity region of the murine NFATp gene. NFAT genes can be used
XX      CC      for prodn. of recombinant NFATs, as probes and primers, and (esp.
XX      CC      antisense) to modulate cellular expression or intracellular
XX      CC      availability of active NFAT.
XX
XX      Sequence 2406 BP; 645 A; 612 C; 535 G; 614 T; 0 other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 2,13e-46 Length: 2406
XX      Score: 430.00 Matches: 79
XX      Percent Similarity: 96.51% Conservative: 4
XX      Best Local Similarity: 91.86% Mismatches: 3
XX      Query Match: 93.68% Indels: 0
XX      DB: 17 Gaps: 0
XX
XX      US-09-550-115A-2 (1-86) x AAT33677 (1-2406)
XX
XX      1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
XX      1171 CTGGCCCTGCATTTTCATTTCACTGCTGTAGACAGTACATCCCTCCCAAAACA 1230
XX
XX      21 ArgLysThrSerLysGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
XX      1231 AGGAAACCTCTAAGATCAAGCTGCCATACACAGGAAATTAAGCTGTTCAGAT 1290
XX
XX      41 AsgGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
XX      1291 GACCAAGGAGGATTATCACCACGCCGCGAGACTTCATATGATATGCGCTTGATCTCAG 1350
XX
XX      61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
XX      1351 TATCTCTTAAGAAAGATTCATGTCGGATCATCTTCTTCACCTTCCTCACCCTTACC 1410
XX
XX      81 TrpSerLysProLysPro 86
XX      1411 TGGAGCAACCAAGCCT 1428
XX
XX      RESULT 2
XX      AAF21183
XX      ID AAF21183 standard; DNA; 3583 BP.
```

```
XX
XX      AAF21183;
AC
XX      14-MAR-2001 (first entry)
XX
XX      Human low adenosine antisense oligonucleotide related sequence #2750.
DT
XX
XX      Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
DE
XX      human; airway disorder; bronchoconstriction; lung inflammation;
DE
XX      surfactant depletion; respiratory; bronchodilator; antiinflammatory;
XX
XX      immunosuppressive; analgesic; hypotensive; cycostatic;
XX      respiratory obstruction; pulmonary vasoconstriction; impeded respiration;
XX      surfactant hypoproduction; pulmonary fibrosis; asthma; RDS;
XX      respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
XX      pulmonary hypertension; emphysema; pulmonary transplantation rejection;
XX      chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
XX      cancer; ss.
XX
XX      Homo sapiens.
OS
XX
XX      WO200062736-A2.
XX
XX      26-OCT-2000.
XX
XX      24-MAR-2000; 2000MO-US08020.
XX
XX      06-APR-1999; 99US-0127958.
XX
XX      (UYEC-) UNIV EAST CAROLINA.
PA
XX      (NYCE/) NYCE J W.
XX
XX      Nyce JW;
XX
XX      WPI: 2000-679539/66.
XX
XX      Low adenosine (A) content antisense oligonucleotides which do not
XX      trigger adenosine receptors during metabolism, useful e.g. for treating
XX      cancers and respiratory obstructions -
XX
XX      PS      Disclosure: Page 1103-1104; 1592pp; English.
XX
XX      The present invention describes low adenosine (A) content antisense
XX      oligonucleotides and compositions (I) comprising them. In the antisense
XX      oligonucleotides the A is replaced by a 'Universal' or alternative base.
XX      (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
XX      immunosuppressive, antisthmatic, hypotensive and cytostatic activities.
XX      The antisense oligonucleotides and (I) can be used to down-regulate the
XX      expression and/or activity of target polypeptides associated with
XX      lung/respiratory disorders and malignancies, such as stimulating and
XX      activating peptide factors and transmitters, transcription factors,
XX      immunoglobulins and antibodies, antibody receptors, cytokines and
XX      chemokines, endogenously produced specific and non-specific enzymes,
XX      binding proteins, adhesion molecules and their receptors, cytokine and
XX      chemokine receptors, adenosine receptors, bradykinin receptors, central
XX      nervous system (CNS), and peripheral nervous and non-nervous system
XX      receptors, CNS and peripheral nervous and non-nervous system peptide
XX      transmitters, defensins, growth factors, vasoactive peptides and
XX      receptors, binding proteins and malignancy associated proteins. The
XX      antisense oligonucleotides may be used in this way to treat disorders
XX      including respiratory obstruction (especially pulmonary obstruction
XX      and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
XX      and/or surfactant hypoproduction which are associated with a disease or
XX      condition selected from pulmonary vasoconstriction, inflammation,
XX      allergies, asthma, impeded respiration, respiratory distress syndrome
XX      (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
XX      hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
XX      pulmonary transplantation rejection, pulmonary infections, bronchitis,
XX      and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
XX      fragments and antisense oligonucleotides used in the exemplification of
XX      the present invention.
XX
XX      Sequence 3583 BP; 959 A; 898 C; 757 G; 969 T; 0 other;
SQ
```


XX	respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KM	pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KM	chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
XX	cancer; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200062736-A2.
XX	
PD	26-OCT-2000.
XX	
PF	24-MAR-2000; 2000WO-US08020.
XX	
PR	06-APR-1999; 99US-0127958.
XX	
PA	(UYEC-) UNIV EAST CAROLINA.
PA	(NICE/) NICE J W.
XX	
PI	Nyce JW;
XX	
DR	WPI; 2000-679539/66.
XX	
PT	Low adenosine (A) content antisense oligonucleotides which do not
PT	trigger adenosine receptors during metabolism, useful e.g. for treating
PT	cancers and respiratory obstructions -
PS	Disclosure; Page 1102-1103; 1592pp; English.
XX	
XX	The present invention describes low adenosine (A) content antisense.
CC	oligonucleotides and compositions (I) comprising them. In the antisense
CC	oligonucleotides the A is replaced by a 'Universal' or alternative base
CC	(I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC	immunosuppressive, antialstmatic, hypotensive and cytostatic activities
CC	The antisense oligonucleotides and (I) can be used to down-regulate the
CC	expression and or activity of target polypeptides associated with
CC	lung/respiratory disorders and malignancies, such as stimulating and
CC	activating peptide factors and transmitters, transcription factors,
CC	immunoglobulins and antibodies, antibody receptors, cytokines and
CC	chemokines, endogenously produced specific and non-specific enzymes,
CC	binding proteins, adhesion molecules and their receptors, cytokine and
CC	chemokine receptors, adenosine receptors, bradykinin receptors, central
CC	nervous system (CNS) and peripheral nervous and non-nervous system
CC	receptors, CNS and peripheral nervous and non-nervous system peptide
CC	transmitters, defensins, growth factors, vasoactive peptides and
CC	receptors, binding proteins and malignancy associated proteins. The
CC	antisense oligonucleotides may be used in this way to treat disorders
CC	including respiratory obstruction (especially pulmonary obstruction
CC	and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC	and/or surfactant hypoproduction which are associated with a disease or
CC	condition selected from pulmonary vasoconstriction, inflammation,
CC	allergies, asthma, impeded respiration, respiratory distress syndrome
CC	(RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC	hyperlension, emphysema, chronic obstructive pulmonary disease (COPD),
CC	pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC	and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC	fragments and antisense oligonucleotides used in the exemplification of
CC	the present invention.
XX	
SO	Sequence 3769 BP; 974 A; 939 C; 802 G; 1054 T; 0 other;
SO	
Alignment Scores:	
Pred. No.:	3,91e-46 Length: 3769
Score:	430.00 Matches: 79
Percent Similarity:	96.51% Conservative: 4
Best Local Similarity:	91.86% Mismatches: 3
Query Match:	93.68% Indels: 0
DB:	Gaps: 0
US-09-555115A-2 (1-86) x AAF21182 (1-3769)	
OY	1 LeuserProAlaProPherProhGlnInfTyrcysValGluThrAspIleProLeuLysThr 20
DB	985 GTTGCCCGCTGCAGTTTTCACATTCACAGACGTGTAGAGACGTGCATCCCTCTCAAAACA 1044

QY	21	ATLVSTHrSerGlaSGLAlAlAlAlLeuProGlyLysLeuGluILeCYsSerAsp	40
QY	1045	AGGAAACTTCTGGAGATCAACAGCTGCCACTACTACCAGGAAATTTAGAGCTGTTCAGAT	1104
QY	41	AspGlnGlyAsnLeuSerProSerArgGluIuHrSerValAspAspGlyLeuGlySerGln	60
Db	1105	GACCAAGGAGTTTATVACCAAGCCCGGAGACTTCATATGATGATGCGCTTGATCTCAG	1164
QY	61	TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr	80
Db	1165	TATCCTTTAAAGAAAGATTCATGTGTGATCATGTTCTTTCAGTTCCCTCACCCCTTAAAC	1224
QY	81	TrpSerLysProLysPro 86	
Db	1225	TGGAGCAACCAACGACCT 1242	
RESULT 5			
AAA35060			
ID	AAA35060	standard; DNA; 3769 BP.	
XX	AAA35060:		
AC	28-JUL-2000	(first entry)	
DT			
XX			
DE	Human adenosine receptor related polynucleotide SEQ ID NO:2749.		
XX			
KW	Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antisthmatic; cytosolic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukemia; lymphoma; carcinoma; metastasis; ss.		
KW			
OS	Homo sapiens.		
XX			
PN	WO200009525-A2.		
XX			
PD	24-FEB-2000.		
XX			
PF	03-AUG-1999; 99WO-US17712.		
XX			
PR	03-AUG-1998; 98US-0095212.		
XX			
PA	(UYEC-) UNITV EAST CAROLINA.		
XX			
PI	Nyce JW;		
DR	WPI: 2000-205971/18.		
XX			
PT	New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers		
PS	Disclosure: Page 1024; 1343pp: English.		
CC	The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytosolic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, sarcomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of		

OY 81 TrpSerLysProLysPro 86
 |||||
 Db 1225 TGGAGCAAAACCAAGCCT 1242
 RESULT 7
 AAA35059
 ID AAA35059 standard; DNA; 3784 BP.
 AC AAA35059;
 XX
 XX 28-JUL-2000 (first entry)
 DT
 XX Human adenosine receptor related polynucleotide SEQ ID NO: 2748.
 DE
 XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytotactic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 KW
 KW Homo sapiens.
 OS
 XX
 XX WO200009525-A2.
 PN
 XX 24-FEB-2000.
 PD
 XX
 XX 03-AUG-1999; 99WO-US17712.
 PF
 XX 03-AUG-1998; 98US-0095212.
 PR
 XX
 XX (UYEC-) UNIV EAST CAROLINA.
 PA
 XX
 XX Nýce JW;
 PI
 XX WPI: 2000-205971/18.
 DR
 XX
 XX New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers
 PS Disclosure; Page 1023; 1343pp: English.
 XX
 XX The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytotactic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasise to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and the then last
 CC 185 sequences are also called SEQ ID NO:1 to 185, and the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA33323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 XX Sequence 3784 BP; 979 A; 950 C; 804 G; 1051 T; 0 other;

Alignment Scores:			
Pred. No.:	3,94e-46	Length:	3784
Score:	430.00	Matches:	79
Percent Similarity:	96.51%	Conservative:	4
Best Local Similarity:	91.86%	Mismatches:	3
Query Match:	93.68%	Indels:	0
DB:	21	Gaps:	0
US-09-550-115A-2 (1-86) x AAA35059 (1-3784)			
QY	1 LeuserProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20		
DB	985 CTGGCCCTCGAGAGTTTTCATTTCATCTGCTGAGAGCTGCATCCCTCTCAACAA 1044		
QY	21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuLysIleCysSerAsp 40		
DB	1045 AGGAAACCTTGTAAGATCAAGCGCCATCTACCGAGAAATTAAGCTGTGTTACAT 1104		
QY	41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60		
DB	1105 GACCAAGGAGATTATTCACACGCGCGGAGACCTCAATAGATGATGGCCTTGATCTCAG 1164		
QY	61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80		
DB	1165 TATCCTTTAAAGAAAGATTCATGTGTGATCATGTTCTTTCAGATTCTTCAACCTTTTACC 1224		
QY	81 TrpSerLysProLysPro 86		
DB	1225 TGGAGCAACCAAGCCCT 1242		
RESULT 8			
ID	ABN95715		
AC	ABN95715 standard; DNA; 3969 BP.		
XX	ABN95715:		
XX	13-AUG-2002 (first entry)		
DE	Gene #2213 used to diagnose liver cancer.		
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;		
KW	metastatic liver tumor; cytostatic; expression profile; disease state;		
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.		
OS	Homo sapiens.		
XX	W0200229103-A2.		
PN	11-APR-2002.		
PD	02-OCT-2001; 2001WO-US30589.		
PF	02-OCT-2000; 2000US-237054P.		
PR	(GENE-) GENE LOGIC INC.		
XX	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;		
PI	WPI; 2002-426119/45.		
XX			
PT	Diagnosing and detecting the progression of liver cancer,		
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,		
PT	involves detecting the level of expression of two or more genes in a		
PT	liver tissue sample		
XX			
PS	Claim 1; SEQ ID NO 2213; 298bp; English.		
CC	The invention relates to a novel method for diagnosing and detecting the		
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver		
CC	tumour in a patient, and differentiating metastatic liver cancer from		
CC	hepatocellular carcinoma in a patient, involving detecting the level of		
CC	expression of two or more genes represented in ABN93503-ABN97455 in a		

sufficient depletion; respiratory bronchiolitis; antiinflammatory;
 immunosuppressive; antisthmatic; analgesic; hypotensive; cytotatic;
 respiratory obstruction; pulmonary obstruction; impeded respiration;
 surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis
 pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 cancer; ss.
 Homo sapiens.
 WO200062736-A2.
 26-OCT-2000.
 24-MAR-2000; 2000WO-US08020.
 06-APR-1999; 99US-0127958.
 (UYEC-) UNIV EAST CAROLINA.
 (NYCE/) NYCE J W.
 Nyce JW;
 WPI; 2000-679539/66.
 Low adenosine (A) content antisense oligonucleotides which do not
 trigger adenosine receptors during metabolism, useful e.g. for treating
 cancers and respiratory obstructions -
 disclosure; Page 1105-1109; 1592pp; English.

CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONS reduces side effects. The A-containing ONS break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA3313 to AAA3512 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 185, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA3323 to AAA3392) are specifically claimed ONS from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 17069 BP: 4137 A; 4938 C; 3964 G; 4030 T; 0 other;

Alignment Scores:
Pred. No.: 3,04e-45 Length: 17069
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
DB: 21 Gaps: 0

US-09-550-115A-2 (1-86) x AAA35063 (1-17069)

OY 1 LeuSerProAlaProPhoPhoPhoGlnItyrCysValGluThrAspIleProLeuLysThr 20
DB 3363 CTGGCCCTGGAGTTTTCATTCAGTGTGAGAGCTGACATCCCTCTCAAAACA 4042

OY 21 ArGlyThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 4043 AGCAAACTTCTGAAGATCAAGCTCCATCTACACGAGAAATTAGACTGTGTCAAT 4102

OY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
DB 4103 GACCAAGCGACTTATATCACAGCCCGGAGACTTCAATATGATGATGAGCTTGAATCTCAG 4162

OY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 4163 TATCTTTAAAGAAAGATTCATGTGTGATCATGTTCTTTCAGTTCCTTACCCCTTACC 4222

OY 81 TrpSerLysProLysPro 86
DB 4223 TCGAGCAACCAACCCCT 4240

RESULT 12
AAC14360/c
ID AAC14360 standard; cDNA: 240 BP.
XX
AC AAC14360:
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 18435.
DE
XX Human: 5' EST: expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX EPI033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-0200610.
PF
XX 26-FEB-1999; 99US-0122487.
PR
XX (GEST) GENSET.
PA
XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
DR
XX
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 18435; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC local human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX
SQ Sequence 240 BP: 81 A; 49 C; 45 G; 59 T; 6 other;

Alignment Scores:
Pred. NO.: 9.79e-16 Length: 240
Score: 187.00 Matches: 34
Percent Similarity: 92.11% Conservative: 1
Best Local Similarity: 89.47% Mismatches: 3
Query Match: 40.74% Indels: 0
DB: 21 Gaps: 0

US-09-550-115A-2 (1-86) x AAC14360 (1-240)

OY 49 ArGluThrSerValAspAspGlyLeuGlySerGlnItyrProLeuLysLysAspSerSer 68
DB 224 CGGAGACTTCAATATGATGATGAGCTGCTGATCTCAKATTCCTTAAAGAAAGATTCATCT 175

OY 69 GlyAspGlnPheLeuSerValProSerProPheThrTrpSerLysProLysPro 86
DB 174 GGTATCAGTTCTTTCAGTTCCTTACCTTACCTTACCTGAGCAACCAACCAACCT 121

RESULT 13
AAT36868
ID AAT36868 standard; cDNA: 2881 BP.
XX
AC AAT36868:
XX
XX 17-NOV-1996 (first entry)
XX
XX Human transcription factor NFAT3 cDNA.
DE
XX Nuclear factor of activated T-cells; NFAT; NFAT3;
KW transcription factor; binding assay; immune system disease;
KW gene therapy; diagnosis; ds.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 142..2850
FT CDS /*tag= a
FT /product= NFAT3
XX
XX W09626959-A1.
PN
XX 06-SEP-1996.
PD
XX 04-MAR-1996; 96WO-US03113.
PF
XX 02-MAR-1995; 95US-0396479.
PR
XX

XX (TULA-) TULARIK INC.
 PA
 XX
 PI Hoey T;
 XX
 DR WPI: 1996-412738/41.
 DR P-PSDB: AAW02250.
 XX
 PT DNA mol. encoding human nuclear factors of activated T cells -
 PT useful for screening potential therapeutic and diagnostic agents for
 PT immune system diseases
 XX
 PS Disclosure: Page 43-47; 64pp; English.
 XX
 CC 3 Overlapping cDNA clones were used to deduce a sequence (AAT36868)
 CC coding for human nuclear factor of activated T-cells class 3. NFAT3
 CC (AAW02250), a regulator of immune system function. The major sites
 CC for NFAT3 expression are outside the immune system, with high
 CC expression in placenta, lung, kidney testis and ovary. NFAT cDNAs
 CC (see also AAT36866-67, AAT3677, AAT36975 and AAT36871) were isolated
 CC from cDNA libraries prep. from Jurkat T-cells and human peripheral
 CC blood lymphocytes by screening with a probe derived from the rel
 CC similarity region of the murine NFATp gene. NFAT genes can be used
 CC for prodn. of recombinant NFATs, as probes and primers, and (esp.
 CC antisense) to modulate cellular expression or intracellular
 CC availability of active NFAT.
 XX
 SQ Sequence 2881 BP; 531 A; 953 C; 882 G; 515 T; 0 other;

Alignment Scores:
 Pred. No.: 1.05e-05 Length: 2881
 Score: 122.00 Matches: 34
 Percent Similarity: 52.94% Conservative: 11
 Best Local Similarity: 40.00% Mismatches: 34
 Query Match: 26.58% Indels: 6
 DB: 17 Gaps: 3

US-09-550-115a-2 (1-86) x AAT36868 (1-2881)

QY 2 SerProAlaProPheProheGlnTrpCysValGluThrAsp---IleProLeuLysThr 20
 DB 1072 TCCCTCGGTCCTTGGACATATGGGGGCGCCACACGACGAGAGCATCCCTCAGAAAGACA 1131
 QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
 DB 1132 CGCGGACTTCCGACGAGCAGCATGCTCTGCTCGCTGAGAGAGCCCTCATGTC 1191
 QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
 DB 1192 AAT---GGGAACTGCTGCTGGAGCAGAGAGAGTGTGCTGCTCCACAGAGGTTCC--- 1245
 QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
 DB 1246 -----CGGAGAGAGTGTGCTGCTGACATGACTGCTGCACTGCCCTCCACATCGCT 1296
 QY 81 TrpSerLysProLys 85
 DB 1297 TGTGTCACAGCCCG 1311

RESULT 14
 AAX08717
 ID AAX08717 standard; DNA; 2881 BP.
 XX
 AC AAX08717;
 XX
 DT 27-SEP-1999 (first entry)
 XX
 DE NF-AT3 encoding sequence.
 XX
 NF-AT3
 KW NF-AT3; hypertrophy; cardiomyocytes; cardiac hypertrophic response;
 KW heart failure; transgenic animals; screening; treatment; inhibition;
 KW ss.
 XX

OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 142..2850
 FT /tag= a
 FT /product= "NF-AT3"
 FT 181..183
 FT /tag= b
 FT /transl_except= AAG encodes Leu
 FT 1126..1128
 FT /tag= c
 FT /transl_except= AAG encodes Leu
 FT 1519..1521
 FT /tag= d
 FT /transl_except= AAG encodes Leu
 FT 1609..1611
 FT /tag= e
 FT /transl_except= AAG encodes Leu
 FT 1654..1656
 FT /tag= f
 FT /transl_except= AAG encodes Leu
 FT 1723..1725
 FT /tag= g
 FT /transl_except= AAG encodes Leu
 FT 1987..1989
 FT /tag= h
 FT /transl_except= AAG encodes Leu
 FT 2020..2022
 FT /tag= i
 FT /transl_except= AAG encodes Leu
 FT 2107..2109
 FT /tag= j
 FT /transl_except= AAG encodes Leu
 FT 2207..2209
 FT /tag= k
 FT /transl_except= AAG encodes Leu
 FT 2653..2657
 FT /tag= l
 FT /transl_except= AAG encodes Leu

W09919471-A1.
 PD 22-APR-1999.
 XX
 PF 15-OCT-1998; 98WO-US21845.
 XX
 PR 16-APR-1998; 98US-0061417.
 PR 16-OCT-1997; 97US-0062864.
 PR 10-NOV-1997; 97US-0065178.
 PR 15-APR-1998; 98US-0081853.
 XX
 PA (UNYNT-) UNIV NORTH TEXAS HEALTH SCI CENT.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Grant SR, Molkenstin JD, Olson EN;
 XX
 DR WPI: 1999-277635/23.
 DR P-PSDB: AAW85730.
 XX
 PT Treating hypertrophy in cardiomyocytes by inhibiting NF-A3
 XX
 PS Disclosure: Page 99-101; 105pp; English.
 XX
 CC Hypertrophy in cardiomyocytes is treated by inhibiting function of
 CC NF-AT3. Activation of NF-AT3 mediates the calcium ion-dependent
 CC cardiac hypertrophic response to a variety of stimuli, so inhibiting
 CC it can be used to treat or prevent cardiac hypertrophy and related
 CC heart failure. Transgenic animals, or cells, containing a
 CC constitutively active NF-AT3 gene can be used as models for screening
 CC modulators of hypertrophy and for studying human disease.
 XX
 SQ Sequence 2881 BP; 531 A; 953 C; 882 G; 515 T; 0 other;

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 18, 2003, 22:06:52 ; Search time 1723 Seconds
(Without alignments) 808.364 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459
Sequence: 1 LSPAPFPQXCVETDPLKT.....SSGDFLVPSPFTWSKPKP 86

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1654066 segs, 809774376 residues

Total number of hits satisfying chosen parameters: 33308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+ p2n.model -DEV=xlh
-O=/cgnr2_1/USPRO.spool/US09550115/unat_11042003_141903_18158/app-query.fasta_1.263
-DB=EST -QFMT=fastap -SUFF1=p2n.rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=XU09550115.GCGN.1.1306.@unat.11042003.141903.18158 -NCPU=6 -ICPU=3
-NO.XLPHY -NO.MAP -LARGOUEVY -NEG.SCORES=0 -MAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hcic:*
9: gb_estcl:*
10: gb_estc2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_plo:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	DB	ID	Description
1	459	100.0	600	12	BC805976	BC805976 2012-19 M
2	459	100.0	600	13	BI990851	BI990851 4082-17 M
3	449	97.8	1073	9	AA821465	AA821465 vs25h04.r
4	430	93.7	455	14	BQ307368	BQ307368 MRO-BT400
5	430	93.7	483	14	BQ307397	BQ307397 MRO-BT400
6	430	93.7	1002	14	BQ070507	BQ070507 AGENCOURT
7	422	91.9	484	14	BQ307403	BQ307403 MRO-BT400
8	418	91.1	502	14	BQ307520	BQ307520 MRO-BT400
9	417	90.8	669	9	AU129466	AU129466 AU129466
10	413	90.0	483	14	BQ307401	BQ307401 MRO-BT400
11	411	89.5	486	14	BQ307399	BQ307399 MRO-BT400
12	410	89.3	920	14	BQ423593	BQ423593 AGENCOURT
13	409	89.1	497	14	BQ307525	BQ307525 MRO-BT400
14	409	89.1	808	13	BI408176	BI408176 602918865
15	402	87.6	379	10	AW244381	AW244381 BR-ENDO3D
16	359	78.2	924	14	BQ214451	BQ214451 AGENCOURT
17	353	76.9	505	14	BQ307486	BQ307486 MRO-BT400
18	328	71.5	501	9	AA840427	AA840427 vx93a04.r
19	316	68.8	1067	13	BI525206	BI525206 602924609
20	313	68.2	728	9	AU123215	AU123215 AU123215
21	259	56.4	490	9	AL120778	AL120778 DKFp762C
22	247	53.8	435	12	BF566087	BF566087 UI-R-BT1
23	209	45.5	918	17	CNS05BC3	ALJ326028 Tetrarodon
24	164.5	35.8	369	9	AA188523	AA188523 zp74c08.r
25	154	33.6	264	13	BJ095964	BJ095964 BI095964
26	145.5	31.7	573	9	AI016619	AI016619 cv30e08.x
27	129	28.1	723	13	BI112028	BI112028 602901368
28	129	28.1	844	11	AK014164	AK014164 Mus muscu
29	117	25.6	1482	13	BM558073	BM558073 AGENCOURT
30	112	25.5	259	9	AA770124	AA770124 bh82c05.s
31	114	24.8	1024	14	BQ710263	BQ710263 AGENCOURT
32	107	23.3	1044	14	BQ714666	BQ714666 AGENCOURT
33	105	22.9	449	12	BF776523	BF776523 287350 MA
34	105	22.9	513	13	BM105571	BM105571 509057 MA
35	103	22.4	461	12	BG746287	BG746287 602703517
36	96.5	21.0	4077	11	AK004810	AK004810 Mus muscu
37	93	20.3	530	14	R87618	R87618 ym91905.r1
38	83	18.1	437	9	AI750778	AI750778 cno5e04.x
39	82	17.9	877	14	BQ733684	BQ733684 AGENCOURT
40	81.5	17.8	594	10	AV827456	AV827456 AV827456
41	81.5	17.8	903	10	BE616886	BE616886 601278756
42	80	17.4	897	9	AL522401	AL522401 AL522401
43	78.5	17.1	1074	17	CNS03LNG	AL246685 Tetrarodon
44	78	17.0	1169	13	BE899396	BE899396 601681370
45	77	16.8	653	12	BM668476	BM668476 UI-E-CL1-

ALIGNMENTS

RESULT 1
BC805976
LOCUS BC805976 600 bp mRNA EST 20-DEC-2001
DEFINITION 2012-19 Mouse E14.5 retina lambda ZAP II library Mus musculus cDNA,
mRNA sequence.
ACCESSION BC805976
VERSION BC805976.1 GI:17952916
KEYWORDS EST.
ORGANISM house mouse.
SOURCE Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murine; Mus.
REFERENCE 1 (bases 1 to 600)
Mu.X., Zhao,S., Pershad,R., Islen,T.-F., Scarpa,A., Wang,S.W.,
White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.

TITLE Gene expression in the developing mouse retina by EST sequencing and microarray analysis
JOURNAL Nucleic Acids Res. 29 (24), 4983-4993 (2001)
MEDLINE 21671825
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329

FEATURES
source Location/Qualifiers
1.600
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone.lib="Mouse E14.5 retina lambda ZAP II Library"
/tissue.type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/note="Vector: PAMPI0 (Gibco): Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Mammatris): Cloning Technique: CVA Cloning (Clontech),
Life Technologies). Average insert size: 1.8 kb;
Insertion site: TACGTCACATGATCTGAGTGC--->. Other
information regarding entire library may be found at
http://pga.smed.edu/Data/Libraries/microarray_cdna_librar
ies.htm."

BASE COUNT 171 a 150 c 121 g 158 t
ORIGIN

Alignment Scores:
Pred. No.: 1.15e-47 Length: 600
Score: 459.00 Matches: 86
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-550-115a-2 (1-86) x BG805976 (1-600)

QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
|||||
DB 35 CTCAGCCCTGCACACCTTTCCATTTCAGTCTGTAGACAGCATCCCTTTGAACA 94
QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerasp 40
|||||
DB 95 AGGAAGACTTCTGACAGATCAAGCTGCCATACACAGAAATTAAGATCTGTTCAGAT 154
QY 41 AspGlnGlyAsnLeuSerProSerArgIuThrSerValAspAspGlyLeuGlySerGln 60
|||||
DB 155 GATCAAGGGAACCTTATCCCTTCCCGGAGACATCATGATGATGATGCGCTTGATCTCAG 214
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
|||||
DB 215 TATCCTTTAAGAAGATTCATCTGTGACCAATTTCTTCACTTCCACCTTTACC 274
QY 81 TrpSerLysProLysPro 86
|||||
DB 275 TGGAGCAAAACCAAGCCT 292

RESULT 2
B1990851 600 bp mRNA linear EST 20-DEC-2001
LOCUS B1990851
DEFINITION 4082-17 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
mRNA sequence.
ACCESSION B1990851
VERSION B1990851.1 GI:17961868
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 600)
Mu X., Zhao S., Pershad R., Hsieh T.-F., Scarpa A., Wang S.W.,
White R.A., Beremand P.D., Thomas T.L., Gan L. and Klein W.H.

TITLE Gene expression in the developing mouse retina by EST sequencing and microarray analysis
JOURNAL Nucleic Acids Res. 29 (24), 4983-4993 (2001)
MEDLINE 21671825
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329

FEATURES
source Location/Qualifiers
1.600
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone.lib="Mouse E14.5 retina lambda ZAP II Library"
/tissue.type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/note="Vector: PAMPI0 (Gibco): Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Mammatris): Cloning Technique: CVA Cloning (Clontech),
Life Technologies). Average insert size: 1.8 kb;
Insertion site: TACGTCACATGATCTGAGTGC--->. Other
information regarding entire library may be found at
http://pga.smed.edu/Data/Libraries/microarray_cdna_librar
ies.htm."

BASE COUNT 158 a 167 c 118 g 157 t
ORIGIN

Alignment Scores:
Pred. No.: 1.15e-47 Length: 600
Score: 459.00 Matches: 86
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-550-115a-2 (1-86) x B1990851 (1-600)

QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
|||||
DB 163 CTCAGCCCTGCACACCTTTCCATTTCAGTCTGTAGACAGCATCCCTTTGAACA 222
QY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerasp 40
|||||
DB 223 AGGAAGACTTCTGACAGATCAAGCTGCCATACACAGAAATTAAGATCTGTTCAGAT 282
QY 41 AspGlnGlyAsnLeuSerProSerArgIuThrSerValAspAspGlyLeuGlySerGln 60
|||||
DB 283 GATCAAGGGAACCTTATCCCTTCCCGGAGACATCATGATGATGATGCGCTTGATCTCAG 342
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
|||||
DB 343 TATCCTTTAAGAAGATTCATCTGTGACCAATTTCTTCACTTCCACCTTTACC 402
QY 81 TrpSerLysProLysPro 86
|||||
DB 403 TGGAGCAAAACCAAGCCT 420

RESULT 3
AA821465/c 1073 bp mRNA linear EST 17-FEB-1998
LOCUS AA821465
DEFINITION vs25h04.r1 Barstead mouse myotubes MFLR5 Mus musculus cDNA clone
IMAGE:1139287 5' similar to gb:U28807 Mus musculus
lymphoid-specific transcription factor NFATc3 (MOUSE);, mRNA
sequence.
ACCESSION AA821465
VERSION AA821465.1 GI:2891333
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1073)

QY 61 TyroProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
 |||
 Db 194 TATCCTTTAAGAAAGATTCATGCTGATCAAGTTCTTCTTCACTTCTTCACTTACC 135
 QY 81 TrpSerLysProLysPro 86
 |||
 Db 134 TGGAGCAACCAAGCCT 117

RESULT 5
 LOCUS BQ307397 483 bp mRNA linear EST 16-MAY-2002
 DEFINITION MR0-BT4000-190601-101-a06 BT4000 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BQ307397
 VERSION BQ307397.1 GI:20843470
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 483)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bais, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL MEDLINE
 COMMENT 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR0&cl2=MR0-BT4000-
 190601-101-a06&cl3=2001-06-19&cl4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 483.
 Location/Qualifiers
 FEATURES
 source
 1..483
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT4000"
 /dev_stage="Adult"
 /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 130 a 102 c 121 g 130 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,85e-44 Length: 483
 Score: 430.00 Matches: 79
 Percent Similarity: 96.51% Conservative: 4
 Best Local Similarity: 91.86% Mismatches: 3
 Query Match: 93.68% Indels: 0
 DB: 14 Gaps: 0
 US-09-550-115A-2 (1-86) x BQ307397 (1-483)
 QY 1 LeuSerProAlaProPheProGlnGlnTyCysValGlnThrAspLeuLysThr 20
 |||
 Db 402 CTTGGCCCTGCAGTTTTCATTTCTGACTGTGTAGAGACATGACATCCCTCAAAACA 343

QY 21 ArgLysThrSerGluAspGlnAlaAlaLeuProGlyLysLeuGlnIleCysSerAsp 40
 |||
 Db 342 AGGAAAACTTGCAGATCAACCTGCCATCACTACCAAGAAATTAGAGCTGTCTCAGAT 283
 QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlnSerGln 60
 |||
 Db 282 GACCAAGGAGGATTATTCACACAGCCCGGAGACTTCAATAGATGATGCTGTGATCTCAG 223
 QY 61 TyroProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
 |||
 Db 222 TATCCTTTAAGAAAGATTCATGCTGATCAAGTTCTTCTTCACTTCTTCACTTACC 163
 QY 81 TrpSerLysProLysPro 86
 |||
 Db 162 TGGAGCAACCAAGCCT 145

RESULT 6
 LOCUS BQ070507 1002 bp mRNA linear EST 02-APR-2002
 DEFINITION AGENCOURT_6839210 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5922897
 5', mRNA sequence.
 ACCESSION BQ070507
 VERSION BQ070507.1 GI:19899553
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1002)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strussberg, Ph.D.
 Email: cgs@bbs-remail.nih.gov
 Tissue procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: L1CM2090 row: 1 column: 10
 High quality sequence stop: 687.
 Location/Qualifiers
 FEATURES
 source
 1..1002
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5922897"
 /clone_lib="NIH_MGC_47"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 264 a 264 c 221 g 251 t 2 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.07e-43 Length: 1002
 Score: 430.00 Matches: 79
 Percent Similarity: 96.51% Conservative: 4
 Best Local Similarity: 91.86% Mismatches: 3
 Query Match: 93.68% Indels: 0
 DB: 14 Gaps: 0
 US-09-550-115A-2 (1-86) x BQ070507 (1-1002)

Alignment Scores:		3,94e-43	Length: 484
Pred. No.:	Score:	422.00	Matches: 78
Percent Similarity:	95.35%	Conservative: 4	
Best Local Similarity:	90.70%	Mismatches: 4	
Query Match:	91.94%	Indels: 0	
DB:	14	Gaps: 0	
US-09-550-115A-2 (1-86) x BQ307403 (1-484)			
QY	1	LeuSerProAlaProPheGlnTyrCysValGluThrAspIleProLeuYThr 20	
Db	403	CTTGCCCTGCAGATTTCATTCATTCAGACTGTGAGAGCATGCATCCCTCAAAACA 344	
QY	21	ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerasp 40	
Db	343	AGGAAACCTTCTGAAGATCAAGCTGCCCTACTACCAGGAAATTAAGACCTGTGCACAT 284	
QY	41	AspIleGlnAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60	
Db	283	GACCAAGGAGATTTATCAACGACCGGAGACCTTCAATAGATGATGCCCTTGATCTGAG 224	
QY	61	TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80	
Db	223	TATCTTTAAAGAGAGATTCATGTGTGATCATCAGTCTCTTCAGTTCACACCTTTACC 164	
QY	81	TrpSerLysProLysPro 86	
Db	163	TGGAGCAAAACCAAGCCCT 146	
RESULT 8	BQ307520	502 bp	mRNA linear EST 16-MAY-2002
LOCUS	BQ307520		
DEFINITION	NR0-BT4000-190601-102-f11 BT4000 Homo sapiens CDNA, mRNA sequence.		
ACCESSION	BQ307520		
VERSION	BQ307520.1	GI:20843835	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 502) Dias Neto,E., Garcia Correa,R., Verjovaki-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags		
JOURNAL MEDLINE COMMENT	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 20202863		
FEATURES	<p>Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br</p> <p>This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=NR0&l2=NR0-BT4000-190601-102-f11&l3=2001-06-19&l4=1)</p> <p>Seq primer: puc 18 forward High quality sequence start: 6 High quality sequence stop: 502.</p> <p>Location/Qualifiers 1..502 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_id="BT4000"</p>		

US-09-550-115A-2 (1-86) x BQ307520 (1-502)

RESULT 9

Eukary

Helix Research Institute

Email: genomics@hri.co.jp

Helix Research Institute.

/organism="Homo sapiens"

Alignment Scores:

US-09-550-115A-2 (1-86) x AU129466 (1-669)

RESULT 10

MEDLINE

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Email: asimpson@ludwig.org.br

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0<2=MR0-BT4000->

190601-101-cl06t3-2001-06-196t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 483.
Location/Qualifiers
1. 483

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BR4000"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 134 a 126 c 97 g 125 t 1 others
ORIGIN

Alignment Scores:
Pred. NO.: 5,37e-42 Length: 483
Score: 413.00 Matches: 76
Percent Similarity: 95.29% Conservative: 5
Best Local Similarity: 89.41% Mismatches: 4
Query Match: 89.98% Indels: 0
DB: 14 Gaps: 0

US-09-550-115A-2 (1-86) x BQ307401 (1-483)

Oy 2 SerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThrArg 21
Db 98 CGCGCTCGAGTTTTCATTTCAGTACTGTAGAGCGCTGACATCCCTCAAAACAGG 157
Oy 22 LysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 41
Db 158 AAAAGCTGTGAGATCAAGCTGCTACTACAGGAAATTAAGACTGTTCACATAC 217
Oy 42 GlnClyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGlnTyr 61
Db 218 CAAGGAGATTATATACGACCGCGGAGACTCAATAGATGATGCGCTTGATCAGATCA 277
Oy 62 ProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThrTrp 81
Db 278 CCTTAAAGAAAGATTCATGTGCTGATCAGTTTCTTTCAGTTCCCTTACCCCTTACCTGG 337
Oy 82 SerLysProLysPro 86
Db 338 AGCAANCAAGCCCT 352

RESULT 11
BQ307399 486 bp mRNA linear EST 16-MAY-2002
LOCUS MRO-BR4000-190601-101-B10 BR4000 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ307399
VERSION BQ307399.1 GI:20843474
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 486)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Goldman, G. H., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 200202653
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MRO-cl2-MRO-BR4000-190601-101-cl06t3-2001-06-196t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 486.
Location/Qualifiers
1. 486

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="BR4000"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 135 a 126 c 100 g 125 t
ORIGIN

Alignment Scores:
Pred. NO.: 9.68e-42 Length: 486
Score: 411.00 Matches: 77
Percent Similarity: 96.47% Conservative: 5
Best Local Similarity: 90.59% Mismatches: 2
Query Match: 89.54% Indels: 1
DB: 14 Gaps: 0

US-09-550-115A-2 (1-86) x BQ307399 (1-486)

Oy 3 ProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThrArgLys 22
Db 101 CCTGAGATTTCCTTCATTTCAGTACTGTAGAGCTGACATCCCTCAAAACAGGAAA 160
Oy 23 ThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAspAspGln 42
Db 161 ACTTCTGAAGATCAAGCTGCCATCTACTACAGGAAATTAAGAGCTGTGATCAGATCACC 220
Oy 43 GlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln-TyrTrp 62
Db 221 GGGAGTTTATCACCGACCGGAGACTTCATAGATGATGCGCTTGATCTCAGTTATCC 280
Oy 62 OleuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThrTrpSe 82
Db 261 TTTAAAGAAAGATTTAGGTGATGATCAGTTTCTTTCAGTTCTTCCCTTACCTTACCTGG 340
Oy 82 LysProLysPro 86
Db 341 CAACCAAGCCCG 353

RESULT 12
BQ423593 920 bp mRNA linear EST 23-MAY-2002
LOCUS AGENCOURT-7762469 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6060527
DEFINITION 5', mRNA sequence.
ACCESSION BQ423593
VERSION BQ423593.1 GI:21118908
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 920)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgi.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1M13349 row: m column: 08
High quality sequence stop: 619.

FEATURES

source

Location/Qualifiers

1..920

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6068527"

/clone_lib="NIH_MGC_92"

/tissue_type="Embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: PCMV-SPORE; site: 1; Note: Site 2: Sail: Cloned unidirectionally; oligo-dt primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

BASE COUNT

222 a 250 c 208 g 240 t

ORIGIN

Alignment Scores:

Pred. No.: 3,17e-41

Score: 410.00

Percent Similarity: 94.25%

Best Local Similarity: 89.66%

Query Match: 89.33%

DB: 14

Matches: 920

Conservative: 78

Mismatches: 4

Indels: 1

Gaps: 0

US-09-550-115a-2 (1-86) x B0423593 (1-920)

OY 1 LeuserProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20

DB 487 CTGGCCCTGCGAGTTTTCATTTTCAGTCTGTAGAGACTGACATCCCTCTCAAAACA 546

OY 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40

DB 547 AGGAAACTTGAAGATCAAGCTGCCATACACAGCAAAATTGAGCTGTTCAGAT 606

OY 41 AspGln-GlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerG 60

DB 607 GACCAAGGCAAGTTTATCCACGCGGAGACTTCAATAGATGATGGCTTGGATCTCA 666

OY 60 nTyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheTh 80

DB 667 GTATCCTTTAAAGAAAGATTGATGATGATGATGATGATGATGATGATGATGATGAT 726

OY 80 rTrpSerLysProLysPro 86

DB 727 CTGGAGCAAAACCAAGCCT 745

RESULT 13 B0307525/c 497 bp mRNA linear EST 16-MAY-2002

LOCUS MR0-BT4000-190601-102-h04 BT4000 Homo sapiens cDNA, mRNA sequence.

DEFINITION B0307525

ACCESSION B0307525.1 GI:20843850

VERSION EST.

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Dias Neto,E., Garcia Correa,R., Veljovski-Almeida,S., Britones,M.R.,

Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordina,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Stimpson,D.H.,

TITLE

JOURNAL

MEDLINE

COMMENT

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR0&l2=MR0-BT4000-

190601-102-h04&l3=2001-06-19&l4=1)

Seq primer: puc 18 forward

High quality sequence start: 16

High quality sequence stop: 497.

Location/Qualifiers

1..497

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="BT4000"

/dev_stage="Adult"

/note="Organ: breast; Vector: puc18; Site: 1; Site: 2: Smat: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196

716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 132 a 108 c 123 g 133 t 1 others

ORIGIN

OY 3 ProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThrArgLys 22

DB 407 CCTGCAAGTTTTCATTTTCAGTCTGTAGAGACTGCATCCCTCGAAACCAAGCAAA 348

OY 23 ThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAspGln 42

DB 347 ACTTCTGAAGATCAAGCTGCCATCTACACGAAATTTAGACTGTGTCAGATGACCGA 288

OY 43 GlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGlnTyrPro 62

DB 287 GGGAGTTTATCANCAAGCCCGGAGACTTCAATAGATGATGGCTTGGATCAGATCTCT 228

OY 63 LeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThrTrpSer 82

DB 227 TTAAGAAAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168

OY 83 LysProLysPro 86

DB 167 AAACCAAGCCT 156

RESULT 14 B1408176/c 808 bp mRNA linear EST 14-AUG-2001

LOCUS B1408176/c

DEFINITION B0291865r1 NCI-GAP-Lu33 Mus musculus cDNA clone IMAGE:5055641 5',

mRNA sequence.

REFERENCE

AUTHORS

```

LOCUS      AM244381/c                      379 bp      mRNA      linear      EST 25-JAN-2002
DEFINITION BR_END03012 Bain Rancourt retinoic acid induced ES cell neural
            differentiation subtraction library Mus musculus cDNA clone 03012
            similar to dbj|D85612|D85612 mouse mRNA for NPATX, complete cds and
            gb|U28807|NMU28807 Mus musculus lymphoidspecific transcription
            factor NFATc3 mRNA, partial cds, mRNA sequence.
ACCESSION  AM244381
VERSION     AM244381.1
KEYWORDS    GI:8051133
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 379)
AUTHORS     Bain,G., Mansergh,F.C., Wide,M.A., Hance,J.E., Isogawa,A.,
            Rancourt,S.L., Ray,W.J., Yoshimura,T., Tsukukl,T., Gottlieb,D.I.
            and Rancourt,D.E.
            ES cell neural differentiation reveals a substantial number of
            novel ESTs
            Funct. Integr. Genomics 1 (2), 127-139 (2000)
            21652683
JOURNAL     Contact: Rancourt DE
MEDLINE     Department of Biochemistry and Molecular Biology
COMMENT     University of Calgary
            3330 Hospital Drive N.W., Calgary, Alberta, T2N 4N1, Canada
            Tel: 403 220 2888
            Fax: 403 283 8727
            Email: rancourt@ucalgary.ca; URL: http://www.acs.ucalgary.ca/(clided
            )rancourt/
            DNA sequencing by: University Core DNA Services, University of
            Calgary. Submitted sequence has been trimmed at both ends to remove
            the adaptor oligos containing the EcoRI sites; i.e. GAATTCGACATCA
            (beginning) and TAGTCCGAATTC (end) removed. Therefore, reported
            insert length is longer than actual EST sequence length.
            Insert length: 389 Std Error: 10.00
            Seq primer: T3 OR T7.
            Location/Qualifiers
                1..379
                   /organism="Mus musculus"
                   /strain="129"
                   /db_xref="taxon:10090"
                   /clone="03D12"
                   /clone_1fb="Bain Rancourt retinoic acid induced ES cell
                   neural differentiation subtraction library"
                   /cell_type="embryonic stem (ES) cell"
                   /cell_line="D3"
                   /lab_host="DH5 alpha"
                   /note="Vector: pBluescript II SK+ (Stratagene); Site_1:
                   EcoRI; Site_2: EcoRI. Library constructed by Dr. Gerard
                   Bain (present address: Hoechst-ARIAD Genomic Center,
                   ARIAD Pharmaceuticals Inc., 26 Landsdowne Street,
                   Cambridge, Massachusetts, 02139-4234, U.S.A.). To isolate
                   cDNAs corresponding to mRNAs which are upregulated during
                   the neural differentiation of ES cells in vitro, the
                   subtractive hybridization technique of Wang and Brown [1]
                   was employed. Poly(A)+ RNA was prepared from both
                   undifferentiated ES cells and from embryoid bodies, which
                   had been cultured for 4 days in the absence of RA followed
                   by an additional 3 days in the presence of 0.5 (M RA
                   (4-/3+ cells). These poly(A)+ RNAs were converted to
                   double-stranded cDNA using the Superscript Choice System
                   (Gibco). Aliquots of both cDNAs were digested with the
                   restriction enzymes AluI and AluI plus PstI. An adaptor
                   oligo [1] containing an EcoRI site was ligated to the ends
                   of the restricted cDNAs to provide primer binding sites
                   and large amounts of each cDNA population were then
                   produced by the polymerase chain reaction (PCR) as
                   described [1]. Amplified cDNA from undifferentiated ES
                   cells was biotinylated using Photoprobe biotin (Vector
                   Laboratories) according to the manufacturer's protocol.
                   2.5 ug of amplified cDNA from 4-/3+ cells was mixed with
                   50 ug of biotinylated ES cell cDNA, denatured by boiling

```

and hybridized for 20 h. Double stranded cDNAs containing biotin were removed by streptavidin/phenol treatment as described [1]. The remaining subtracted cDNA was mixed with an additional 25 mg of biotinylated ES cell cDNA, denatured by boiling, and hybridized for 2 h. The streptavidin/phenol treatment was repeated and the remaining cDNA was amplified by PCR [Wang and Brown, 1991]. Two additional rounds of subtraction were repeated exactly as described above. The cDNA obtained from this subtraction procedure was digested with EcoRI and ligated to pBS II SK+ (Stratagene) followed by transformation into E. coli DH5 cells. Individual colonies were picked and the corresponding plasmids were isolated either by an alkaline lysis miniprep procedure [2], or using the Qiaprep spin miniprep kit (Qiagen). Sequence analysis was performed using the Big Dye Cycle Sequencing kit and an ABI373 sequencer (University Core DNA Services, University of Calgary). 1. Wang, Z.; Brown, D.D. (1991) A gene expression screen. Proc. Natl. Acad. Sc 1. USA 88, 11505-11509. 2. Sambrook, J.; Fritsch, E.F.; Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

BASE COUNT 124 a 65 c 95 g 95 t
ORIGIN

Alignment Scores:

Pred. No.:	9.31e-41	Length:	379
Score:	402.00	Matches:	76
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	87.58%	Indels:	0
DB:	10	Gaps:	0

US-09-550-115A-2 (1-86) x AW244381 (1-379)

```

QY 11 CysValGluThrAspIleProLeuLysThrArgLysThrSerGluAspGlnAlaIle 30
DB 376 TGTGTAGAGACTGACATCCCTTGAAACACAGAAAGACTTGAAGATCAAGCTGCCATA 317
QY 31 LeuProGlyLysLeuGluIleCysSerAspAspGlnGlyAsnLeuSerProSerArgGlu 50
DB 316 CTACCGAGAAATTAAGATCTCTTCAGATGATCAAGGAAGACTTATCCCTCCGGGAG 257
QY 51 ThrSerValAspAspGlyLeuGlySerGlnTyProLeuLysLysAspSerSerGlyAsp 70
DB 256 ACATCAGTAGATGATGGCTTGATCTCAGTATCCTTTAAAGAAAGATCATCTGGTGAC 197
QY 71 GlnPheLeuSerValProSerProPheThrTrpSerLysProLysPro 86
DB 196 CAATTCTTTCAGTCTTCACCTTACCTGGAGCAACCAAGCCT 149

```

Search completed: April 18, 2003, 23:53:27
Job time : 1734 secs

GenCore version 5.1.4-P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 18, 2003, 22:21:47 ; Search time 53 seconds
(without alignments)
497.626 Million cell updates/sec

Title: US-09-550-115a-2
Perfect score: 459
Sequence: 1 LSPAPPPFQYCVETDIPDKT.....SSGDFLSVSPFTSKKPP 86

Scoring table:
BLOSUM62
Xgapop 10.0 , Ygapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=x1h
-O=/cgn2_1/USPTO.spool/US09550115/rnat_11042003_141904_18174/app_query.fasta.1.263
-DB=Issued_Patents_NA -OPMT=fasta -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEXT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09550115.ecgn.1.1.40 @unat.11042003.141904.18174 -NCPU=6 -ICPU=3
-NO.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/line/5A.COMB.seq.*
3: /cgn2_6/ptodata/1/line/5B.COMB.seq.*
4: /cgn2_6/ptodata/1/line/6A.COMB.seq.*
5: /cgn2_6/ptodata/1/line/6B.COMB.seq.*
6: /cgn2_6/ptodata/1/line/backfillseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	430	93.7	2406	1 US-08-396-479B-7	Sequence 7, Appl1
2	430	93.7	2406	1 US-08-818-823-7	Sequence 7, Appl1
3	430	93.7	2647	1 US-08-396-479B-9	Sequence 9, Appl1
4	430	93.7	2647	1 US-08-818-823-9	Sequence 9, Appl1
5	430	93.7	3969	1 US-08-396-479B-11	Sequence 11, Appl1
6	430	93.7	3969	1 US-08-818-823-11	Sequence 11, Appl1
7	430	93.7	4010	5 PCT-US94-07297-40	Sequence 40, Appl1
8	122	26.6	2881	1 US-08-396-479B-5	Sequence 5, Appl1
9	122	26.6	2881	1 US-08-818-823-5	Sequence 5, Appl1
10	114	24.8	2743	1 US-08-396-479B-3	Sequence 3, Appl1
11	114	24.8	2743	1 US-08-818-823-3	Sequence 3, Appl1
12	114	24.8	2749	2 US-08-124-981A-1	Sequence 1, Appl1

13	114	24.8	2751	3 US-09-037-190-45	Sequence 45, Appl1
14	114	24.8	2751	3 US-09-037-192-45	Sequence 45, Appl1
15	114	24.8	2751	3 US-09-037-143-45	Sequence 45, Appl1
16	114	24.8	2751	4 US-09-049-691-45	Sequence 45, Appl1
17	114	24.8	2751	4 US-08-260-174-45	Sequence 45, Appl1
18	114	24.8	2751	4 US-09-338-128A-45	Sequence 45, Appl1
19	114	24.8	2751	4 US-09-037-192-45	Sequence 45, Appl1
20	114	24.8	2853	5 PCT-US94-07297-36	Sequence 36, Appl1
21	114	24.8	5502	4 US-09-232-346-45	Sequence 45, Appl1
22	71.5	15.6	2340	3 US-09-022-883-4	Sequence 4, Appl1
23	71.5	15.6	2477	4 US-09-490-692-3	Sequence 3, Appl1
24	70	15.3	2019	2 US-08-245-511-46	Sequence 46, Appl1
25	70	15.3	2019	2 US-08-600-993A-46	Sequence 46, Appl1
26	70	15.3	11864	4 US-08-961-527-61	Sequence 61, Appl1
27	67.5	14.7	1386	2 US-08-408-095-22	Sequence 22, Appl1
28	67	14.6	4527	3 US-09-045-632-1	Sequence 1, Appl1
29	67	14.6	6506	4 US-09-453-702B-1	Sequence 1, Appl1
30	65.5	14.3	6360	4 US-09-171-699-9	Sequence 9, Appl1
31	64	13.9	1266	4 US-08-790-186A-3	Sequence 3, Appl1
32	64	13.9	1445	4 US-08-949-246-1	Sequence 1, Appl1
33	63	13.7	1289	4 US-09-149-476-319	Sequence 319, Appl1
34	62.5	13.6	576	1 US-08-086-428B-48	Sequence 48, Appl1
35	62.5	13.6	576	2 US-08-468-570-48	Sequence 48, Appl1
36	62.5	13.6	576	2 US-08-290-665A-48	Sequence 48, Appl1
37	62.5	13.6	576	5 PCT-US95-10398-48	Sequence 48, Appl1
38	62.5	13.6	1302	4 US-09-255-368-5	Sequence 5, Appl1
39	62.5	13.6	4339	4 US-09-484-970B-164	Sequence 164, Appl1
40	62.5	13.6	5077	2 US-08-687-956A-22	Sequence 22, Appl1
41	62.5	13.6	49136	4 US-09-422-869-1	Sequence 1, Appl1
42	62	13.5	628	4 US-08-998-416-87	Sequence 87, Appl1
43	62	13.5	1567	3 US-08-961-083-119	Sequence 119, Appl1
44	62	13.5	2557	4 US-08-464-954A-1	Sequence 1, Appl1
45	62	13.5	3121	3 US-08-961-083-117	Sequence 117, Appl1

ALIGNMENTS

RESULT 1
US-08-396-479B-7
Sequence 7, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLEHR, HOERBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8771
TELEFAX: (415) 494-8700
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2406 base pairs
Type: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..2337
US-08-396-479B-7

Alignment Scores:
Pred. No.: 5.5e-48 Length: 2406
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-396-479B-7 (1-2406)

OY 1 LeuserProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
DB 1171 CTTGGCCCTGCGAGTTTTCATTTCAGTGTGTAGAGACTGACATCCCTCTCAAAACA 1230

OY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 1231 AGGAAACTTCTGAAGATCAAGCTCCACTACTACCAAGAAATTAGAGCTGTTCAGAT 1290

OY 41 AspGlnGlyAsnLeuSerProSerArgLysThrSerValAspAspGlyLeuGlySerGln 60
DB 1291 GACCAAGGAGATTATCAACACCGCGGAGACTTCATATAGATGATGCGCTTGATCTCAG 1350

OY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 1351 TATCCTTTAAAGAAAGATCATGTGTGATCAGTTCTTTCAGTTCTTCACCTTTAC 1410

OY 81 TrpSerLysProLysPro 86
DB 1411 TGGACCAACCAACCT 1428

RESULT 2
US-08-818-823-7
Sequence 7, Application US/08818823
Patent No. 5708158
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,823
FILING DATE: 14-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/396,479
FILING DATE: 02-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2406 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..2337
US-08-818-823-7

Alignment Scores:
Pred. No.: 5.5e-48 Length: 2406
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-818-823-7 (1-2406)

OY 1 LeuserProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
DB 1171 CTTGGCCCTGCGAGTTTTCATTTCAGTGTGTAGAGACTGACATCCCTCTCAAAACA 1230

OY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 1231 AGGAAACTTCTGAAGATCAAGCTCCACTACTACCAAGAAATTAGACTGTTCAGAT 1290

OY 41 AspGlnGlyAsnLeuSerProSerArgLysThrSerValAspAspGlyLeuGlySerGln 60
DB 1291 GACCAAGGAGATTATCAACACCGCGGAGACTTCATATAGATGATGCGCTTGATCTCAG 1350

OY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 1351 TATCCTTTAAAGAAAGATCATGTGTGATCAGTTCTTTCAGTTCTTCACCTTTAC 1410

OY 81 TrpSerLysProLysPro 86
DB 1411 TGGACCAACCAACCT 1428

RESULT 3
US-08-396-479B-9
Sequence 9, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771

TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 09:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..2427
US-08-396-479B-9

Alignment Scores:
Pred. No.: 6,26e-48 Length: 2647
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-396-479B-9 (1-2647)

OY 1 LeuSerProAlaProPheProPheGlnIrrCysValGluThrAspIleProLeuLysThr 20
DB 1171 CTTGGCCCTGCGAGTTTTCATTTCAGTCTGTAGAGACTGACATCCCTCTCAAAACA 1230

OY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 1231 AGCAAACTCTCGAAGATCAACCTCCATCAACGAAATTTAGAGCTGTCTCAGAT 1290

OY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
DB 1291 GACCAAGGAGATTATTCACACCGCGAGACTTCATATGATGATGAGCCCTTGATCTCAG 1350

OY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 1351 TATCCTTTAAAGAAATTCATGCTGATCAGTTCTTTCAGTTCCCTTACCCCTTTACC 1410

OY 81 TrpSerLysProLysPro 86
DB 1411 TCGACCAAAACCAAGCCT 1428

RESULT 4
US-08-818-823-9
Sequence 9, Application US/08818823
Patent No. 5708158
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,823
FILING DATE: 14-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/396,479
FILING DATE: 02-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 09:
SEQUENCE CHARACTERISTICS:
LENGTH: 2647 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..2427
US-08-818-823-9

Alignment Scores:
Pred. No.: 6,26e-48 Length: 2647
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-818-823-9 (1-2647)

OY 1 LeuSerProAlaProPheProPheGlnIrrCysValGluThrAspIleProLeuLysThr 20
DB 1171 CTTGGCCCTGCGAGTTTTCATTTCAGTCTGTAGAGACTGACATCCCTCTCAAAACA 1230

OY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
DB 1231 AGCAAACTCTCGAAGATCAACCTCCATCAACGAAATTTAGAGCTGTCTCAGAT 1290

OY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
DB 1291 GACCAAGGAGATTATTCACACCGCGAGACTTCATATGATGATGAGCCCTTGATCTCAG 1350

OY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
DB 1351 TATCCTTTAAAGAAATTCATGCTGATCAGTTCTTTCAGTTCCCTTACCCCTTTACC 1410

OY 81 TrpSerLysProLysPro 86
DB 1411 TCGACCAAAACCAAGCCT 1428

RESULT 5
US-08-396-479B-11
Sequence 11, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A

```
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..3414
US-08-396-479B-11

Alignment Scores:
Pred. No.: 1.09e-47 Length: 3969
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-396-479B-11 (1-3969)

QY 1 LeuSerProAlaProPheProPheGlnIrrCysValGluThrAspIleProLeuIsthr 20
Db 1171 CTTGGCCCTGCACTTTTTCATTTCAGTCTGTAGAGACTGACATCCCTCCAAACA 1230
QY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
Db 1231 AGGAAACTCTCAATACAGCTGCCATACATACAGAAATTAAGCTGTGTACAGT 1290
QY 41 AspGlnGlyAsnLeuSerProSerArgIuThrSerValAspAspGlyLeuGlySerGln 60
Db 1291 GACCAAGGAGATTATACACAGCCGCGAGACTTCAATAGATGATGCTTGATCTCAG 1350
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1351 TATCCTTTAAGAAATTCATGTGTGATCAGTTCTTTCAGTTCTTCACCTTTACC 1410
QY 81 TrpSerLysProLysPro 86
Db 1411 TGGAGCAACCAAGCCT 1428

RESULT 6
US-08-818-823-11
Sequence 11, Application US/08818823
Patent No. 5708158
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,823
FILING DATE: 14-MAR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: 08/396,479
FILING DATE: 02-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 3969 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 211..3414
US-08-818-823-11

Alignment Scores:
Pred. No.: 1.09e-47 Length: 3969
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
Gaps: 0
DB: 1

US-09-550-115A-2 (1-86) x US-08-818-823-11 (1-3969)

QY 1 LeuSerProAlaProPheProPheGlnIrrCysValGluThrAspIleProLeuIsthr 20
Db 1171 CTTGGCCCTGCACTTTTTCATTTCAGTCTGTAGAGACTGACATCCCTCCAAACA 1230
QY 21 ArgLysThrSerGluAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
Db 1231 AGGAAACTCTCAATACAGCTGCCATACATACAGAAATTAAGCTGTGTACAGT 1290
QY 41 AspGlnGlyAsnLeuSerProSerArgIuThrSerValAspAspGlyLeuGlySerGln 60
Db 1291 GACCAAGGAGATTATACACAGCCGCGAGACTTCAATAGATGATGAGCTTGATCTCAG 1350
QY 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1351 TATCCTTTAAGAAATTCATGTGTGATCAGTTCTTTCAGTTCTTCACCTTTACC 1410
QY 81 TrpSerLysProLysPro 86
Db 1411 TGGAGCAACCAAGCCT 1428

RESULT 7
PCT-US94-07297-40
Sequence 40, Application PC/TUS9407297
GENERAL INFORMATION:
APPLICANT: Arai, Naoko
APPLICANT: Masuda, Etsuo
APPLICANT: Tokumitsu, Hiroshi
TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN
TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: John H. C. Blasdale, Schering-Plough Corporation, M-3-W
STREET: One Giralda Farms
CITY: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 07940-1000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: System Software 7.1
```

SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07297
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/222,626
FILING DATE: 04-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,061
FILING DATE: 05-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,971
FILING DATE: 30-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,998
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/088,483
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFERENCE/DOCKET NUMBER: DX0392K4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 4010 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 304..3531
FEATURE:
NAME/KEY: unsure
LOCATION: replace(1756..1758, "gta")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(1756..1758, "gaa")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(1756..1758, "gga")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(1756..1758, "gga")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(3090..3092, "agt")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(3090..3092, "aga")
FEATURE:
NAME/KEY: unsure
LOCATION: replace(3090..3092, "agg")
PCT-US94-07297-40
Alignment Scores:
Pred. No.: 1,1e-47 Length: 4010
Score: 430.00 Matches: 79
Percent Similarity: 96.51% Conservative: 4
Best Local Similarity: 91.86% Mismatches: 3
Query Match: 93.68% Indels: 0
DB: 5 Gaps: 0
US-09-550-115a-2 (1-86) x PCT-US94-07297-40 (1-4010)
Oy 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
Db 1264 CTGGCCCTGCGATTTTCATTCAGTGTGAGAGACTGACATCCCTCAAAACA 1323
Oy 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
|||||

Db 1324 AGAAACTTCTGAAGATCAAGCTGCCACTACTACAGGAAATTAAGAGCTGTTCAGAT 1383
Oy 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
Db 1384 GACCAAGGAGATTATACACGAGCCCGAGACTTCAATGATGAGCCCTGGATCTCAG 1443
Oy 61 TyrProLeuLysLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1444 TATCCTTAAAGAAAGATTCATGTGTGATGATCAGTTTCTTTCAGTTCCTTCAACCTTTAAC 1503
Oy 81 TrpSerLysProLysPro 86
Db 1504 TGGAGCAAAACCAAGCCT 1521
RESULT 8
US-08-396-479B-5
Sequence 5, Application US/08396479B
Patent No. 5612455
GENERAL INFORMATION:
APPLICANT: HOEY, Timothy
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOENACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,479B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59450-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 210 277299
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2881 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2850
US-08-396-479B-5
Alignment Scores:
Pred. No.: 1,31e-06 Length: 2881
Score: 122.00 Matches: 34
Percent Similarity: 52.94% Conservative: 11
Best Local Similarity: 40.00% Mismatches: 34
Query Match: 26.58% Indels: 6
DB: 1 Gaps: 3
US-09-550-115a-2 (1-86) x US-08-396-479B-5 (1-2881)
Oy 2 SerProAlaProPheProPheGlnTyrCysValGluThrAspIleProLeuLysThr 20
Db 1072 TCCCTGCTGCTTTCATATGCTGGGGCCCGACAGCTGAGAGATCCCTCAGAAAGCA 1131
Oy 21 ArgLysThrSerGluAspGlnAlaAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
|||||

Db 1132 CGCGGACTTCCAGGAGAGGAGGAGTGGCTGCTCGCTCGGTCTGAGAGGAGCTTCATGC 1191
QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
Db 1192 AAT---GGGAGACTCCCTTGGGAGACAGAGAGTGTGTGCTCTCCAGAGGTTCC--- 1245
QY 61 TyrProLeuLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1246 -----CGGAGAGAGGTGGCTGGCATGACTGACTGCGTCCCTCCACTCGCT 1296
QY 81 TrpSerLysProLys 85
Db 1297 TGGTCCAAAGCCCCG 1311
RESULT 9
US-08-818-823-5
; Sequence 5, Application US/08818823
; Patent No. 5708158
; GENERAL INFORMATION:
; APPLICANT: HOEY, Timothy
; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,823
; FILING DATE: 14-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/396,479
; FILING DATE: 02-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59450-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 210 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2881 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2850
US-08-818-823-5
Alignment Scores:
Pred. No.: 1,31e-06 Length: 2881
Score: 122.00 Matches: 34
Percent Similarity: 52.94% Conservative: 11
Best Local Similarity: 40.00% Mismatches: 34
Query Match: 26.58% Indels: 6
DB: 1 Gaps: 3
US-09-550-115A-2 (1-86) x US-08-818-823-5 (1-2881)
QY 2 SerProAlaProPheProPheGlnTyrCysValGluThrAsp---IleProLeuLysThr 20

Db 1072 TCCCTGTGCTCCCTTGGACTATGTGGGGCCGCCACAGCTGAGAGCATCCCTCAGAAAGACA 1131
QY 21 ArgLysThrSerGlnAspGlnAlaIleLeuProGlyLysLeuGluIleCysSerAsp 40
Db 1132 CGCGGACTTCCAGGAGAGGAGGAGTGGCTGCTCGGTCTGAGAGGAGCTTCATGC 1191
QY 41 AspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAspGlyLeuGlySerGln 60
Db 1192 AAT---GGGAGACTCCCTTGGGAGACAGAGAGTGTGTGCTCTCCAGAGGTTCC--- 1245
QY 61 TyrProLeuLysAspSerSerGlyAspGlnPheLeuSerValProSerProPheThr 80
Db 1246 -----CGGAGAGAGGTGGCTGGCATGACTGACTGCGTCCCTCCACTCGCT 1296
QY 81 TrpSerLysProLys 85
Db 1297 TGGTCCAAAGCCCCG 1311

RESULT 10
US-08-396-479B-3
; Sequence 3, Application US/08396479B
; Patent No. 5612455
; GENERAL INFORMATION:
; APPLICANT: HOEY, Timothy
; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,479B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59450-1/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 210 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2743 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 240..2390
US-08-396-479B-3
Alignment Scores:
Pred. No.: 1,45e-05 Length: 2743
Score: 114.00 Matches: 25
Percent Similarity: 61.11% Conservative: 19
Best Local Similarity: 34.72% Mismatches: 26
Query Match: 24.84% Indels: 2
DB: 1 Gaps: 2
US-09-550-115A-2 (1-86) x US-08-396-479B-3 (1-2743)

[illegible]

```

Oy      36  GlyIecysSerAspApgInGlyAsnLeuSerProSerArgGIuThrValaspSsp 55
        |||.....|||.....|||...:::...:::||||
Db      1288  GAGCCCGTGGGAGAGACCTGGGCAGCCCCGCCGCCGCGACTTGCGCCCGCAAAAC 1347
        |||.....|||.....|||...:::...:::||||

Oy      56  GlyLeuGIySerGIuTyrrProleuLytyAspSerSerGIyAspGlnPheLeuSerVal 75
        |||:::.....:::|||||
Db      1348  TACHTCCTTTCCAGCAC--ATCAGGAAGGGCGGCTTGTGGACACAGTAAGTGGCGGTG 1404
        |||.....|||:::.....|||.....|||.....|||

Oy      76  ProSer---ProPhenTrTPSerLysPrIoLySpro 86
        |||.....|||:::.....|||.....|||.....|||
Db      1405  CCGCAGCACCCCTACAGTAGTGGCGGAGCCCAAGCCCAAGCCC 1440

RESULT 13
US-09-037-190-45
; Sequence 45, Application US/09037190
; Patent No. 6096515
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: No. 6096515throp, Jeffrey P.
; APPLICANT: Ho, Steffan M.
; TITLE OF INVENTION: NF-A1 POLYPEPTIDES AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,190
FILING DATE: 09-MAR-1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,174
FILING DATE: 13-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/124,981
FILING DATE: 20-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-332.03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 240..2387
US-09-037-190-45

Alignment Scores:
Pred. No.: 1.45e-05 Length: 2751
Score: 114.00 Matches: 25
Percent Similarity: 61.1% Conservative: 19
Best Local Similarity: 34.72% Mismatches: 26
Query Match: 24.84% Indels: 2
DB: 3 Gaps: 2

US-09-550-115A-2 (1-86) x US-09-037-190-45 (1-2751)
```

```

Oy 16 IleProLeuSerThrArgGlySerGlnAlaIleLeuProGlyLeuSer 35
Db 1230 GTCCCTGTCAAGTCCGCCAGACACCCTTGAGACAGACCTCAAGTGGCCTCAAGGTG 1289
Oy 36 GluIleCysSerAspAspGlnGlyAsnLeuSerProSerArgGluThrSerValAspAsp 55
Db 1290 GAGCCGCGGAGGAGGAGCTGCGACGCCGCCGCCGCCGCCGCCGCCGCCGAGAC 1349
Oy 56 GlyLeuGlySerGlnGlyrrProLeuGlySlyAspSerSerGlyAspGlnPheLeuSerAl 75
Db 1350 TACTCTCTTTCAGCAC--ATCAGAGAAAGGGCGGCTTCTGCGACCACTGACTGGCGGTG 1406
Oy 76 ProSer---ProPheThrTPSerTyrProTyrPro 86
Db 1407 CCGCAGCACCCCTACACAGTGGGCGGAGGCCACGCC 1442

RESULT 14
US-09-037-192-45
Sequence 45, Application US/09037192
Patent No. 6096860
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: No. 6096860throp, Jeffrey P.
APPLICANT: Ho, Steffan M.
TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,192
FILING DATE: 09-MAR-1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,174
FILING DATE: 13-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/124,981
FILING DATE: 20-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-332.04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEO ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 2751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 240..2387
US-09-037-192-45

Alignment Scores:
Pred. No.: 1,45e-05 Length: 2751
Score: 114.00 Matches: 25
Percent Similarity: 61.118 Conservative: 19
Best Local Similarity: 34.728 Mismatches: 26

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: April 18, 2003, 22:25:43 ; Search time 118 Seconds

(Without alignments)
732.138 Million cell updates/sec

Title: US-09-550-115A-2

Perfect score: 459
Sequence: 1 LSPAPPEFFQYCVETDIPLKT.....SSGDQLSVSPFTWSKPKP 86

Scoring table:

BLSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=Frame+p2n.model -DEV=xlh
-O/cgn2.1/USPRO.spool/US09550115/runat.11042003.141905.18204/app.query.fasta.1.263
-DB=Published.Applications_NA -QFW=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dlsum62
-TRANS=naun40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09550115 -ECGN.1.1.93 -runat.11042003.141905.18204
-NCPU=6 -ICPU=3 -NO.XLPXY -NO.MMAP -LARGQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications_NA:*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PC7_NEM_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	430	93.7	3969	10 US-09-880-107-2213	Sequence 2213, Ap
2	114	24.8	5502	9 US-10-040-430-45	Sequence 45, Appl
3	114	24.8	5502	9 US-09-349-058-45	Sequence 45, Appl
4	76	16.6	32768	10 US-09-070-927A-17	Sequence 17, Appl

5	71	15.5	32189	9 US-10-079-854-379	Sequence 379, App
6	71	15.5	32189	10 US-09-764-878-379	Sequence 379, App
7	71	15.5	32221	9 US-10-079-854-377	Sequence 377, App
8	71	15.5	32221	10 US-09-764-878-377	Sequence 377, App
9	69	15.0	7132	12 US-10-044-090-780	Sequence 780, App
10	67.5	14.7	379	9 US-10-046-935-435	Sequence 435, App
11	67.5	14.7	379	9 US-09-878-178-435	Sequence 435, App
12	67.5	14.7	379	9 US-10-146-502-435	Sequence 435, App
13	67	14.6	6506	9 US-10-114-170-1	Sequence 1, Appl1
14	66	14.4	463	10 US-09-244-694-87	Sequence 87, Appl1
15	66	14.4	500	10 US-09-864-761-12394	Sequence 12394, A
16	66	14.4	503	9 US-09-736-457-1364	Sequence 1364, Ap
17	66	14.4	503	9 US-09-902-941-1364	Sequence 1364, Ap
18	66	14.4	503	9 US-09-849-626-1364	Sequence 1364, Ap
19	66	14.4	503	9 US-10-017-754-1364	Sequence 1364, Ap
20	66	14.4	1214	10 US-09-764-864-87	Sequence 87, Appl
21	65.5	14.3	2567	9 US-09-925-299-167	Sequence 167, App
22	65.5	14.3	2567	10 US-09-925-299-167	Sequence 167, App
23	64.5	14.1	530	9 US-09-796-692-9499	Sequence 9499, Ap
24	64.5	14.1	1887	10 US-09-954-456-1624	Sequence 1624, Ap
25	64.5	14.1	2085	9 US-09-405-920-1	Sequence 1, Appl1
26	64.5	14.1	6663	10 US-09-037-657-28	Sequence 28, Appl
27	64.5	14.1	11832	10 US-09-037-657-38	Sequence 38, Appl
28	64	13.9	1634	9 US-10-108-605-30	Sequence 30, Appl
29	64	13.9	2076	10 US-09-823-847-1	Sequence 1, Appl1
30	64	13.9	6339	10 US-09-727-770-3	Sequence 3, Appl1
31	63	13.7	550	10 US-09-560-863-463	Sequence 463, App
32	63	13.7	766	10 US-09-910-943-489	Sequence 489, App
33	63	13.7	1023	10 US-09-822-849A-537	Sequence 537, App
34	63	13.7	1289	9 US-09-809-351-319	Sequence 319, App
35	63	13.7	1817	10 US-09-764-864-457	Sequence 457, App
36	63	13.7	32195	9 US-10-092-154-1512	Sequence 1512, Ap
37	63	13.7	32195	10 US-09-764-847-1512	Sequence 1512, Ap
38	62.5	13.6	400	7 US-08-781-986A-1158	Sequence 1158, Ap
39	62.5	13.6	942	9 US-09-938-842A-1789	Sequence 1789, Ap
40	62.5	13.6	1289	9 US-09-988-442-55	Sequence 55, Appl
41	62.5	13.6	1289	9 US-10-073-865-50	Sequence 50, Appl
42	62.5	13.6	1289	10 US-09-764-853-240	Sequence 240, App
43	62.5	13.6	1302	9 US-09-866-248A-5	Sequence 5, Appl1
44	62.5	13.6	1739	10 US-09-731-872-225	Sequence 225, App
45	62.5	13.6	49136	10 US-09-768-877-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-880-107-2213
Sequence 2213, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880.107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2213
LENGTH: 3969
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L41067

Alignment Scores:


```

Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunach
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 32768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-070-927A-17

Alignment Scores:
Pred. NO.: 15.4 Length: 32768
Score: 76.00 Matches: 23
Percent Similarity: 48.28% Conservative: 19
Best Local Similarity: 26.44% Mismatches: 27
Query Match: 16.56% Indels: 18
DB: Gaps: 4

US-09-550-115A-2 (1-86) x US-09-070-927A-17 (1-32768)
OY 3 ProAlaProPherProPhelInTrCySValGIuThAspLleProLeuLysThrArgLys 22
Db 65 CCGAACACTTACACGCGTCAAAWTTGATGTAGAAACCGAGGCTTTTAAATCAAAAAA 124
OY 23 ThrSerGIuAspGlnAlaIleAlaIleLeuProGlyLysLeuGluIleCysSerAspGln 42
Db 125 GTTGATTAAGATCAAGCTGCTATTGTACACAGA-----ACGGTTTCCATTAGACTTT 178
OY 43 GlyAsnLeuSerProSerArgLuhTrSerValAsp-----Asp 55
Db 179 GGAAAAAAGCTTACCGCAAAAGACGCTGACTGACAAAGAAAGCATTCGTACATTGAT 238
OY 56 GlyLeu-----GlySerClnTrProLeuLysLysAspSerSerGlyAspGlnPheLeu 73
Db 239 GCGATTCTCATGCTGACAAAGGTAACCATTTACCGAAAA----- 277
OY 74 SerValProSerProPhetheThr 80

```

```

Db      278  TCGCTGCCGGCCCTTATACG  298
|||||
RESULT 5
US-10-079-854-379
; Sequence 379, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 379
; LENGTH: 32189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-079-854-379

Alignment Scores:
Pred. No.:      76.6      Length:      32189
Score:          71.00     Matches:      19
Percent Similarity: 52.83% Conservative: 9
Best Local Similarity: 35.85% Mismatches: 15
Query Match:    15,47%   Indels:      10
DB:              9      Gaps:        3

US-09-550-115A-2 (1-86) x US-10-079-854-379 (1-32189)

QY      32  PROGLYLVLSLEUGLIILECYSSerSaspSgInGLyAsnLeuSerProSerArgcIuNrHr  51
|||||
Db      25457  CCAGGAGAGGCCATTGTGTGTCTTGTGGGAGCAGCTGCCGAGAGCTCCCGAGGAGAA  2516
|||||
QY      52  SerValAspSgIlyLeuGlySerGlnTrpProLeuLysAspSerSerGlyAspGln  71
|||||
Db      25517  GCA-----GAAGGAGCTTGAGGCCACAGCAGCCT-----GGTCTACT  25522
|||||
QY      72  PheLeuSerValProSerProPhe-----ThrTrpSer  82
|||||
Db      25553  CCTCTGACATCCAGCCACTTTTCTCTGCACTTGTCG  25591
|||||

RESULT 6
US-09-764-878-379
; Sequence 379, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 379
; LENGTH: 32189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-379

Alignment Scores:
Pred. No.:      76.6      Length:      32189
Score:          71.00     Matches:      19
Percent Similarity: 52.83% Conservative: 9
Best Local Similarity: 35.85% Mismatches: 15
Query Match:    15,47%   Indels:      10
DB:              10     Gaps:        3

US-09-550-115A-2 (1-86) x US-09-764-878-379 (1-32189)

```

```

Oy 32 ProGlyLysLeuGluIleCySerAspArgInGlyAsnLeuSerProSerArgGluThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25457 CCAGGAGAGGCACTTGTGTGTCTGCGGAGCAGCTCCGAGAGCTCCCGAGGGAAGAA 25516
Oy 52 SerValAspAspGlyLeuGlySerGlnIleProLeuLysAspSerSerGlyAspGln 71
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25517 GCA-----GAAGAGCACTTGAGGCGCCAGACACT-----GGTCTAGT 25552
Oy 72 PheLeuSerValProSerProPhe-----ThrTrpSer 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25553 CCTCTGACATCCCGACCACTTTTCTGCTGCGACTTGCTCC 25591

RESULT 7
US-10-079-854-377
: Sequence 377, Application US/10079854
: Publication No. US20030054368A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA121
: CURRENT APPLICATION NUMBER: US/10/079,854
: CURRENT FILING DATE: 2002-02-22
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 428
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 377
: LENGTH: 32221
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (7464)
: OTHER INFORMATION: n equals a,t,g, or c
US-10-079-854-377

Alignment Scores:
Pred. NO.: 76.7 Length: 32221
Score: 71.00 Matches: 19
Percent Similarity: 52.83% Conservative: 9
Best Local Similarity: 35.85% Mismatches: 15
Query Match: 15.47% Indels: 10
DB: 9 Gaps: 3

US-09-550-115A-2 (1-86) x US-10-079-854-377 (1-32221)
Oy 32 ProGlyLysLeuGluIleCySerAspArgInGlyAsnLeuSerProSerArgGluThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25488 CCAGGAGAGGCACTTGTGTGTCTGCGGAGCAGCTCCGAGAGCTCCCGAGGGAAGAA 25547
Oy 52 SerValAspAspGlyLeuGlySerGlnIleProLeuLysAspSerSerGlyAspGln 71
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25548 GCA-----GAAGAGCACTTGAGGCGCCAGACACT-----GGTCTAGT 25583
Oy 72 PheLeuSerValProSerProPhe-----ThrTrpSer 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25584 CCTCTGACATCCCGACCACTTTTCTGCTGCGACTTGCTCC 25622

RESULT 8
US-09-764-878-377
: Sequence 377, Application US/09764878
: Patent No. US20020090615A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PA121
: CURRENT APPLICATION NUMBER: US/09/764,878
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 428
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 377
: LENGTH: 32221
: TYPE: DNA

```

```

: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (7464)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-878-377

Alignment Scores:
Pred. No.:          76-7       Length:      32221
Score:              71.00     Matches:      19
Percent Similarity: 52.83%    Conservative: 9
Best Local Similarity: 35.85% Mismatches:    15
Query Match:        15.47%   Indels:      10
DB:                  10      Gaps:         3

US-09-550-115A-2 (1-86) x US-09-764-878-377 (1-32221)
OY      32 ProglyLysLeuGluIleCySerAspSpdInGlyAsnLeuSerProserArgIuThr 51
           |||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 25488 CCAGCAAGCGCATTTGTCTGTTCGGGGACAGCTGCAGGACTCCACAGAGGAAGA 25547
           ::::::::::::::::::::|||:::|||||:::|||||

OY      52 SerValaspaspgLyLeuGLySerGIuTyrProLeuLysLysaspserGIyaspGln 71
           GCA-----GAGAGACTTGAGGCCCAAGCACC-----GCTGCTAGT 25583
           ::::::::::::::::::::|||:::|||||:::|||||

Db 25548 GCA-----GAGAGACTTGAGGCCCAAGCACC-----GCTGCTAGT 25583
           ::::::::::::::::::::|||:::|||||:::|||||

RESULT 9
US-10-044-090-780/c
: Sequence 780, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olgja Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-0028 US
: CURRENT APPLICATION NUMBER: US/10/044,090
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO 780
: LENGTH: 7132
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20020137081A1 198499.13
US-10-044-090-780

Alignment Scores:
Pred. No.:          21         Length:      7132
Score:              69.00     Matches:      13
Percent Similarity: 51.22%    Conservative: 8
Best Local Similarity: 31.71% Mismatches:     6
Query Match:        15.03%   Indels:      14
DB:                  12      Gaps:         1

US-09-550-115A-2 (1-86) x US-10-044-090-780 (1-7132)
OY      7 ProphegIntyCyValaGluThrAspIleProLeuLysThrArgLysThrSerGIuaSp 26
           |||  |||  |||||:::|||||:::|||||:::|||||
Db 3561 CCAACCCAGGCTGTGTATGAGAGACCTCCGCTCAGAAC----- 3520
           ::::|||||:::|||||:::|||||:::|||||

OY      27 GlinaAlaAlaIleuProGlyLysLeuGluIleCySeraspSpdInGlyasnLeuSer 46
           ::::|||||:::|||||:::|||||:::|||||
Db 3519 -----ATCCAAGTGTGCCCGTTCGATCTTGCGCAATGTGGCC 3484
           ::::|||||:::|||||:::|||||:::|||||

RESULT 10
```


Db 256 GATGTCCTGAGTTTACGACCGCTTTGTGCGGGAGTGAGGCTGGGGGTG 197
QY 60 GlnTyrProLeuLysAspSerSerGlyAspGlnPheLeuSerValProSer---Pro 78
Db 196 AGAAAGGAGAGAGGGGAGGAGGACAGCTTCTTTTCTCTAGCTTACC 137
QY 79 PheThrTrpSerLysProLys 85
Db 136 TTTTCTTAATAGCCCAA 116
RESULT 13
US-10-114-170-1/c
Sequence 1, Application US/10114170
Publication No. US20030023075A1
GENERAL INFORMATION:
APPLICANT: Blatner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Plinkney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296, 95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6506
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-114-170-1
Alignment Scores:
Pred. No.: 35.7 Length: 6506
Score: 67.00 Matches: 27
Percent Similarity: 36.36% Conservative: 9
Best Local Similarity: 27.27% Mismatches: 23
Query Match: 14.60% Indels: 40
Gaps: 4
US-09-550-115A-2 (1-86) x US-10-114-170-1 (1-6506)
QY 18 LeuLysThrArgLysThrSerGlnuSpGlnAlaIleLeuProGlyLysLeuGluIle 37
Db 4352 TTAATAACCGTTCAACCTCATGAC-----TTAGAAATT 4317

QY 38 CysSerAspAspGlnGlyAsnLeuSerProSerArgLysThr----- 51
Db 4316 CAGTGTGAC-----GGAATATACAGCCCAACTATGCAAGTACTAGCATCAGATTGAT 4263
QY 52 -----SerVal 53
Db 4262 GCTAATACTTACGCTGATTAAGAACTGCTGAGAACATTGGCTTGCATTTCCTC 4203
QY 54 AspAspGlyLeu-----GlySerGlnTyrProLeuLysAspSer 67
Db 4202 GATGATGGAGCTTTGACCAAGATGACACCGGTTAAATAACACCGCTGAAGATATATGGC 4143
QY 68 SerGlyAspGlnPheLeuSerValProSerProPheThrTrpSerLysProLysPro 86
Db 4142 TCTGTCATATTTTCTCTCATTTTAATGCAATATGCCCGGACAAATGCTACT 4086
RESULT 14
US-09-244-694-87
Sequence 87, Application US/09244694
Patent No. US20020026037A1
GENERAL INFORMATION:
APPLICANT: Hu, Jing-Shan
APPLICANT: Olsen, Henrik S.
APPLICANT: Rosen, Craig A.
TITLE OF INVENTION: Vascular Endothelial Growth Factor 3
FILE REFERENCE: 1488.1040003
CURRENT APPLICATION NUMBER: US/09/244,694
CURRENT FILING DATE: 1999-02-10
EARLIER APPLICATION NUMBER: US 09/132,088
EARLIER FILING DATE: 1998-08-10
EARLIER APPLICATION NUMBER: US 09/033,662
EARLIER FILING DATE: 1998-03-03
EARLIER APPLICATION NUMBER: US 08/469,641
EARLIER FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 87
LENGTH: 463
TYPE: DNA
ORGANISM: Mus musculus
US-09-244-694-87
Alignment Scores:
Pred. No.: 1.63 Length: 463
Score: 66.00 Matches: 28
Percent Similarity: 43.12% Conservative: 19
Best Local Similarity: 25.69% Mismatches: 34
Query Match: 14.38% Indels: 29
Gaps: 5
US-09-550-115A-2 (1-86) x US-09-244-694-87 (1-463)
QY 1 LeuSerProAlaProPheProPheGlnTyrCysValGluThrAspLysProLeuLysThr 20
Db 67 ATCTCCCGAGTCAAACTCCACCTGCGACCTCTCCCGGAGGGGTGCGCCGCGACG 126
QY 21 Arg-----LysThrSerGluAspGlnAlaIleLeuProGlyLys 34
Db 127 CCGGAGCGACCGCTTGCGCCCAAGAACGAGACCTCGGGTGTCCCGCGCTACCCG 186
QY 35 LeuGluIleCysSerAspAspGlnGlyAsnLeuSerPro---SerArgLysThrSerVal 53
Db 187 GTCAGTGA-CTCAAGATGAGTAGTGTATTTCACCGCGCGCCGCGAGACCGCGG 245
QY 54 -----AspAspGlyLeuGlySerGlnTyrPro----- 62
Db 246 CCGGACCCCGACGAGAGACGCGCGCGCGCGCGCTCACACTTATTCTACACT 305
QY 63 -----LeuLysLysAspSerSerGlyAspGlnPheLeuSer 74
Db 306 CTCATGTCTTTCACCGTGCAGACTAGAGTCAAGCTCAACAGGGGTCTTTCCTCCCGCT 365

PI Hoey T;
 XX MPI: 1996-412738/41.
 DR N-PSDB; AAT33677.
 XX
 PT DNA mol. encoding human nuclear factors of activated T cells -
 PT useful for screening potential therapeutic and diagnostic agents for
 PT immune system diseases
 PS Claim 6; Page 52-54; 64pp; English.
 XX
 CC 4 Types of human nuclear factor of activated T-cells class 4,
 CC NFAT4a (AAW02251), NFAT4b and NFAT4c, result from alternative
 CC splicing downstream of the rel homology domain. The 3 types
 CC have identical N-terminal sequences, but C-terminal sequences
 CC differ for NFATb (AAW02252) and NFATc (AAW02253) from that for
 CC NFATa. NFATs (see also AAW02248-50) include regulators of cytokine
 CC gene expression that modulate immune system function. Recombinant
 CC NFATs, or NFAT fragments contg. the rel domain, can be expressed in
 CC prokaryotic or eukaryotic host cells. They are used in high-
 CC throughput screenings to identify agents useful in the diagnosis or
 CC treatment of diseases associated with expression of a gene modulated
 CC by a transcription complex contg. NFAT(s).
 XX
 SQ Sequence 708 AA;
 Query Match 93.7%; Score 430; DB 17; Length 708;
 Best Local Similarity 91.9%; Pred. No. 3.2e-46;
 Matches 79; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LSPAPPFQYCVETDIPKTRKTSDDAAILPGKLEICSDDGNSPSRRTSVDDGLGSG 60
 DB 321 LSPAVPPFQYCVETDIPKTRKTSDDAAILPGKLEICSDDGSLSPARETSIDDELGSG 380
 QY 61 YPLKDSGDOFLSVSPPTWSKPKP 86
 DB 381 YPLKDSGDOFLSVSPPTWSKPKP 406
 RESULT 2
 AAR75353
 ID AAR75353 standard; Protein; 1075 AA.
 XX
 AC AAR75353;
 DT 12-AUG-1995 (first entry)
 XX
 DE Human NF-AT120, X subfamily.
 XX
 KW NF-AT120; nuclear factor of activated T cells 120 protein;
 KW transcription; differentiation; cytokine; T-cell; T-lymphocyte.
 OS Homo sapiens.
 XX
 PN W09502053-A.
 PD 19-JAN-1995.
 PE 05-JUL-1994; 94WO-US07297.
 PF 06-JUL-1993; 93US-0088483.
 PR 30-JUL-1993; 93US-0099998.
 PR 30-AUG-1993; 93US-0113971.
 PR 05-NOV-1993; 93US-0148061.
 PR 04-APR-1994; 94US-0222626.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Arai N, Masuda ES, Tokumitsu H;
 CC MPI: 1995-066898/09.
 DR N-PSDB; AAW084303.
 XX

PT Nuclear Factor of Activated T cells 120 protein, antibody and
 PT nucleic acid - useful as transcriptional regulator of genes, such
 PT as cytokines
 XX
 PS Disclosure; Page 78-83; 100pp; English.
 XX
 CC A DNA fragment amplified from Jurkat CDNA by PCR had the sequence
 CC given in AAW084300 and encoded NF-AT120 (Nuclear Factor of
 CC Activated T cells protein 120) (AAR66873). Amplified DNA was used
 CC to screen for related NF-AT proteins by hybridization, leading to
 CC the discovery of 3 related subfamilies, designated class C, P and
 CC X (AAR66874, AAR66875, AAR75353).
 XX
 SQ Sequence 1075 AA;
 Query Match 93.7%; Score 430; DB 16; Length 1075;
 Best Local Similarity 91.9%; Pred. No. 5.6e-46;
 Matches 79; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LSPAPPFQYCVETDIPKTRKTSDDAAILPGKLEICSDDGNSPSRRTSVDDGLGSG 60
 DB 321 LSPAVPPFQYCVETDIPKTRKTSDDAAILPGKLEICSDDGSLSPARETSIDDELGSG 380
 QY 61 YPLKDSGDOFLSVSPPTWSKPKP 86
 DB 381 YPLKDSGDOFLSVSPPTWSKPKP 406
 RESULT 3
 AAW02250
 ID AAW02250 standard; Protein; 902 AA.
 XX
 AC AAW02250;
 DT 17-NOV-1996 (first entry)
 XX
 DE Human transcription factor NFAT3.
 XX
 KW Nuclear factor of activated T-cells; NFAT; NFAT3;
 KW transcription factor; cytokine; gene expression; binding assay;
 KW immune system disease; therapy; diagnosis.
 OS Homo sapiens.
 XX
 PN W09626959-A1.
 PD 06-SEP-1996.
 PE 04-MAR-1996; 96WO-US03113.
 PR 02-MAR-1995; 95US-0396479.
 PA (TULA-) TULARIK INC.
 PI
 PI Hoey T;
 XX
 DR MPI: 1996-412738/41.
 DR N-PSDB; AAT36868.
 PT DNA mol. encoding human nuclear factors of activated T cells -
 PT useful for screening potential therapeutic and diagnostic agents for
 PT immune system diseases
 PS Claim 5; Page 43-47; 64pp; English.
 XX
 CC The amino acid sequence (AAW02250) of human nuclear factor of activated
 CC T-cells class 3, NFAT3, was deduced from an isolated cDNA clone
 CC (AAT36868). NFATs (see also AAW02248-49 and AAW02251-53) include
 CC regulators of cytokine gene expression that modulate immune system

[illegible]


```

XX  MO9626959-A1.
XX  06-SEP-1996.
XX  04-MAR-1996; 96WO-US03113.
XX  02-MAR-1995; 95US-0396479.
XX  (TULA-) TULARIK INC.
XX  Hoey T.
XX  WPI: 1996-412738/41.
XX  DR  N-PSDB; AAT36867.
XX  PT  DNA mol. encoding human nuclear factors of activated T cells -
XX  useful for screening potential therapeutic and diagnostic agents for
XX  immune system diseases
XX  PS  Claim 4; Page 41-43; 64pp; English.
XX  CC  The amino acid sequence (AAW02249) of human nuclear factor of activated
XX  T-cells class C, NFATC, was deduced from an isolated cDNA clone
XX  (AAT36867). NFATs (see also AAW02249 and AAW02250-53) include
XX  regulators of cytokine gene expression that modulate immune system
XX  function. They have invariant rel domain peptides (see also AAW02254-55)
XX  and share at least 50% sequence identity in their rel domains.
XX  CC  Recombinant NFATs, or NFAT fragments contg. at least part of the rel
XX  domain, can be expressed in prokaryotic or eukaryotic host cells.
XX  CC  They are used in high-throughput screenings to identify agents useful
XX  in the diagnosis or treatment of diseases associated with expression
XX  of a gene modulated by a transcription complex contg. NFAT(s).
XX  SO  Sequence 716 AA:

Query Match 24.8%; Score 114; DB 17; Length 716;
Best Local Similarity 34.7%; Pred. No. 1.6e-05;
Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps 2;

OY  16 IPLKTRKTSDDOAILPKLEICSDGDNLSPSRSTVSDDGSGYPLKDSGQPLSV 75
DB  331 VPVSKRKTTLQPPSVALKVERVEGDLGSPPPADFAEDYSSFOH-IRKGGFCQYLAIV 389
OY  76 PS-PPTWSKRP 86
DB  390 PQHRYQMAKRP 401

RESULT 9
AAV96535
ID  AAV96535 standard; Protein: 716 AA.
XX  AAV96535:
XX  12-SEP-2000 (first entry)
XX  Human nuclear factor of activated T cells c1 (NF-ATc1).
XX  NF-ATc1: cardiac hypertrophy; nuclear factor of activated T cells;
XX  antagonist; congestive heart disease; cardiact.
XX  Homo sapiens.
XX  OS  MO200030671-A2.
XX  PN  02-JUN-2000.
XX  PD  23-NOV-1999; 99WO-US27862.
XX  PF  24-NOV-1998; 98US-0198977.
XX  PR  (STRD ) UNIV LELAND STANFORD JUNIOR.

```

```

XX  Crabtree GR, Northrop JP, Ho SN;
XX  WPI: 2000-399929/34.
XX  DR  N-PSDB; AAA29252.
XX  PT  Treating cardiac hypertrophy using NF-AT antagonists in
XX  patients suffering from congestive heart disease
XX  PS  Disclosure; Fig 1A-B; 139pp; English.
XX  CC  A novel method for preventing and/or reducing cardiac hypertrophy in a
XX  patient, comprises administering an NF-AT (nuclear factor of activated
XX  T cells) antagonist to decrease the biological activity of NF-AT in
XX  myocardial tissue (therefore preventing and/or reducing the level of
XX  cardiac hypertrophy). The antagonists may decrease the transcriptional
XX  activity, nuclear translocation or dephosphorylation of NF-AT, inhibit
XX  binding of calcineurin to NF-AT, stimulate phosphorylation of NF-AT (e.g.
XX  by increasing GSK-3) or inhibit formation of an NF-AT complex. The
XX  CC  antagonist is an antagonist of NF-ATc4 (also known as NF-AT3) and not
XX  NF-ATc1, NF-ATc2 and NF-ATc3 (also known as NF-AT4). The method is used
XX  CC  for preventing and/or reducing cardiac hypertrophy in a patient suffering
XX  CC  from congestive heart disease (claimed) and for preventing other growth
XX  CC  of cardiac and vascular tissue.
XX  SO  Sequence 716 AA:

Query Match 24.8%; Score 114; DB 21; Length 716;
Best Local Similarity 34.7%; Pred. No. 1.6e-05;
Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps 2;

OY  16 IPLKTRKTSDDOAILPKLEICSDGDNLSPSRSTVSDDGSGYPLKDSGQPLSV 75
DB  331 VPVSKRKTTLQPPSVALKVERVEGDLGSPPPADFAEDYSSFOH-IRKGGFCQYLAIV 389
OY  76 PS-PPTWSKRP 86
DB  390 PQHRYQMAKRP 401

RESULT 10
AAV96565
ID  AAV96565 standard; Protein: 716 AA.
XX  AAV96565:
XX  12-SEP-2000 (first entry)
XX  Human nuclear factor of activated T cells c1 (NF-ATc1).
XX  NF-ATc1: cardiac hypertrophy; nuclear factor of activated T cells;
XX  KW  antagonist; congestive heart disease; cardiact.
XX  OS  Homo sapiens.
XX  FH  Key Location/Qualifiers
XX  FT  Modified-site 172..301
XX  FT  Domain 172..194
XX  FT  /label= SRR
XX  FT  /note= "NLS interacting domain"
XX  FT  199..219
XX  FT  Domain 199..219
XX  FT  /label= SPI
XX  FT  /note= "NLS interacting domain"
XX  FT  233..252
XX  FT  /label= SP2
XX  FT  /note= "NLS interacting domain"
XX  FT  265..267
XX  FT  Misc-difference 265..267
XX  FT  /label= "Nuclear localization_signal"
XX  FT  278..301
XX  FT  Domain 278..301
XX  FT  /label= SP3
XX  FT  /note= "NLS interacting domain"

```

FT	Misc-difference	681..685
FT	/label=	Nuclear_Localization_signal
FT	/note=	"C-terminal"
XX		
PN	WO200030671-A2.	
XX		
PD	02-JUN-2000.	
XX		
PF	23-NOV-1999;	99WO-US27862.
XX		
PR	24-NOV-1998;	98US-0198977.
XX		
PA	(STRD) UNIV LELAND STANFORD JUNIOR.	
PI	Crabtree GR, Northrop JP, Ho SN;	
DR	WPI: 2000-399929/34.	
XX		
PT	Treating cardiac hypertrophy using NF-AT antagonists in	
PT	patients suffering from congestive heart disease	
XX		
PS	Disclosure; Page 49; 139pp; English.	
XX		
CC	A novel method for preventing and/or reducing cardiac hypertrophy in a	
CC	patient, comprises administering an NF-AT (nuclear factor of activated	
CC	T cells) antagonist to decrease the biological activity of NF-AT in	
CC	myocardial tissue (therefore preventing and/or reducing the level of	
CC	cardiac hypertrophy). The antagonists may decrease the transcriptional	
CC	activity, nuclear translocation or dephosphorylation of NF-AT, inhibit	
CC	binding of calcineurin to NF-AT, stimulate phosphorylation of NF-AT (e.g.	
CC	by increasing GSK-3) or inhibit formation of an NF-AT complex. The	
CC	antagonist is an antagonist of NF-ATc4 (also known as NF-ATc3) and not	
CC	NF-ATc1, NF-ATc2 and NF-ATc3 (also known as NF-ATd). The method is used	
CC	for preventing and/or reducing cardiac hypertrophy in a patient suffering	
CC	from congestive heart disease (claimed) and for preventing other growth	
CC	of cardiac and vascular tissue.	
XX		
SO	Sequence	716 AA:
	Query Match	24.8%; Score 114; DB 21; Length 716;
	Best Local Similarity	34.7%; Pred.No. 1.6e-05;
	Matches	25; Conservative 19; Mismatches 26; Indels 2; Gaps 2
QY	16 IPLKRRKSSEDAALLPKGLICSDDGRLSRETSDVDGSLQPLPKDDSSGGDFLSV	75
	:1:1:1:1:1: :	
DB	331 VPVKSRKRTLLDPSPVALKVPEPVGEDLSPPPADFAPEDYSSFQH-IRKGGFCDGYIAV	389
	:	
QY	76 PS-PFTWSKP RP 86	
	: : :	
DB	390 PQHRYQMAKRPK 401	
	RESULT 11	
	AAB90774	
ID	AAB90774 standard; Protein; 716 AA.	
XX		
AC	AAB90774;	
XX		
DJ	15-JUN-2001 (first entry)	
XX		
DE	Human shear stress-response protein SEQ ID NO: 48.	
XX		
KM	Human; Shear stress-response protein; vascular disease;	
XX	arteriosclerosis.	
OS	Homo sapiens.	
PN	MO200125427-A1.	
XX		
FD	12-APR-2001.	
XX		
FE	02-OCT-2000; 2000WO-JP06840.	
XX		

PR	01-OCT-1999;	99JP-0280976.	
XX			
PA	(KYOW) KYOWA HAKKO KOGYO KK.		
PA	(NOJI/) NOJIMA H.		
XX			
PI	Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;		
PI	Kuga T, Sekine S, Nakamura Y, Sugano S;		
XX			
DR	WPI: 2001-266308/27.		
DR	N-PSDB: AAH02897.		
XX			
PT	DNA sequences, proteins encoded by them and antibodies against them		
PT	useful in diagnosis and treatment of vascular disease caused by		
PT	arteriosclerosis -		
XX			
PS	Claim 60; Page 352-356; 678pp; Japanese.		
XX			
CC	The present invention provides the protein and coding sequences of a		
CC	number of human shear stress response proteins. These are useful in th		
CC	diagnosis, treatment and screening of vascular diseases caused by		
CC	arteriosclerosis, including heart failure, post-PTCA restenosis and		
XX	hypertension.		
SO	Sequence 716 AA;		
	Query Match	24.88;	Score 114; DB 22; Length 716;
	Best Local Similarity	34.7%;	Pred. No. 1.6e-05;
	Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps		
QY	16	ILPLTKRTSEDOAAILPLCKLEICSDDOGNLSPSRERISVDDGLGSGYPRKKDSSGDPFLSY 75	
DB	331	VPVASKRTTLEQPPSVALKVEPVGEDLGSPPPADPAEDYSSFOH-IRKGGFCDOYLAV 389	
QY	76	PS-PFTWSKPKP 86	
DB	390	PQHPEYQWAKPKP 401	
	RESULT 12		
	AAB66497		
ID	AAB66497	standard; Protein; 716 AA.	
XX			
AC	AAB66497;		
XX			
DT	10-APR-2001	(first entry)	
XX			
DE	Human NF-AT protein.		
XX			
KW	Human; nuclear factor of activated T lymphocytes; NF-AT;		
KW	NF-AT cytoplasmic component; NF-ATc; NF-AT translocation;		
KW	nuclear localisation sequence; NLS; SRR.		
XX			
OS	Homo sapiens.		
XX			
NN	US6171781-BL.		
XX			
PD	09-JAN-2001.		
XX			
PF	27-MAR-1998;	98US-0049691.	
XX			
XX	20-SEP-1993;	93US-0124981.	
PR	13-JUN-1994;	94US-0260174.	
XX			
PA	(STRD) UNIV LELAND STANFORD JUNIOR.		
PI	Crabtree GR, Northrop JP, Ho SN;		
XX			
DR	WPI; 2001-122328/13.		
XX			
PT	Screening assay for identifying modulators of translocation of nuclear		
PT	factor of activated T lymphocytes across nuclear membrane of cell, by		
DT	treating with nuclear factor of activated T lymphocyte polypeptide		

PS Claim 10; Column 40; 99pp; English.
XX
CC The present sequence is given in a specification relating to a method for
CC identifying a compound which modulates translocation of a nuclear factor
CC of activated T lymphocytes (NF-AT) polypeptide across the nuclear
CC membrane of a cell. The method involves binding the compound to the
CC NF-AT polypeptide. The method is useful for identifying compounds which
CC modulate nuclear translocation of NF-AT present in the cytoplasm and/or
CC in the nucleus of the cell and for identifying agents that modulate
CC phosphorylation/dephosphorylation of NF-AT.
XX
SQ Sequence 716 AA;
Query Match 24.8%; Score 114; DB 22; Length 716;
Best Local Similarity 34.7%; Pred. No. 1.6e-05;
Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps 2;
OY 16 IPLKTRKTSSEOAILPKLEICSDOGNLSPSRSTSVDDGLGSGYPLKDKSSGDOFLSV 75
DB 331 VPVKSRRKTTLEQPSVALKVEPVGEDLGSPPPADFAFEDYSSFOH-IRKGGFCDOYLAV 389
OY 76 PS-PPTMSKPKP 86
DB 390 PQHPYOMAKPKP 401
RESULT 13
AAB6501
ID AAB6501 standard; Protein: 716 AA.
XX
AC AAB6501;
XX
DT 10-APR-2001 (first entry)
XX
DE Human NF-ATc protein.
XX
KW Human; nuclear factor of activated T lymphocytes; NF-AT;
KW NF-AT cytoplasmic component; NF-ATc; NF-AT translocation;
KW nuclear localisation sequence; NLS; SRR.
XX
OS Homo sapiens.
XX
PA US6171781-B1.
XX
PI Crabtree GR, Northrop JP, Ho SN;
XX
DR WPI: 2001-122328/13.
DR N-PSDB: AAF31710.
XX
PT Screening assay for identifying modulators of translocation of nuclear
PT factor of activated T lymphocytes across nuclear membrane of cell, by
PT treating with nuclear factor of activated T lymphocyte polypeptide -
XX
PS Disclosure: Fig 1; 99pp; English.
XX
CC The present sequence is given in a specification relating to a method for
CC identifying a compound which modulates translocation of a nuclear factor
CC of activated T lymphocytes (NF-AT) polypeptide across the nuclear
CC membrane of a cell. The method involves binding the compound to the
CC NF-AT polypeptide. The method is useful for identifying compounds which
CC modulate nuclear translocation of NF-AT present in the cytoplasm and/or
CC in the nucleus of the cell and for identifying agents that modulate
CC phosphorylation/dephosphorylation of NF-AT.

SQ Sequence 716 AA;
Query Match 24.8%; Score 114; DB 22; Length 716;
Best Local Similarity 34.7%; Pred. No. 1.6e-05;
Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps 2;
OY 16 IPLKTRKTSSEOAILPKLEICSDOGNLSPSRSTSVDDGLGSGYPLKDKSSGDOFLSV 75
DB 331 VPVKSRRKTTLEQPSVALKVEPVGEDLGSPPPADFAFEDYSSFOH-IRKGGFCDOYLAV 389
OY 76 PS-PPTMSKPKP 86
DB 390 PQHPYOMAKPKP 401
RESULT 14
AAE21559
ID AAE21559 standard; Protein: 716 AA.
XX
AC AAE21559;
XX
DT 16-JUL-2002 (first entry)
XX
DE Human cytoplasmic nuclear factor of activated T cell (NF-ATc) protein #1.
XX
KW Human; immunosuppressive; cytoplasmic nuclear factor of activated T cell;
KW NF-ATc; nuclear translocation.
XX
OS Homo sapiens.
XX
PN US6352830-B1.
XX
PD 05-MAR-2002.
XX
PE 15-JAN-1999; 99US-0223246.
XX
PR 22-AUG-1991; 91US-0749385.
PR 18-APR-1994; 94US-0228944.
PR 20-SEP-1993; 93US-0124981.
PR 13-JUN-1994; 94US-0260174.
PR 31-JUL-1995; 95US-0507032.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Crabtree GR, Northrop JP, Ho SN, Flanagan WM;
XX
DR WPI: 2002-314700/35.
XX
PT Identifying immunosuppressive agent comprises contacting cell having
PT cytoplasmic NF-AT polypeptide with inducer of polypeptide cytoplasmic
PT translocation, in presence and absence of test agent, and assaying the
PT translocation -
XX
PS Example 9; Column 67-72; 83pp; English.
XX
CC The invention relates to a method for identifying an immunosuppressive
CC agent. The method comprising: contacting a cell containing cytoplasmic
CC nuclear factor of activated T cell (NF-ATc) polypeptide with a compound
CC that induces nuclear translocation of the polypeptide; and nuclear
CC translocation of the NF-ATc is assayed. The method is useful for
CC identifying an immunosuppressive agent and an immune regulating agent.
CC The present sequence is human NF-ATc protein.
XX
SQ Sequence 716 AA;
Query Match 24.8%; Score 114; DB 22; Length 716;
Best Local Similarity 34.7%; Pred. No. 1.6e-05;
Matches 25; Conservative 19; Mismatches 26; Indels 2; Gaps 2;
OY 16 IPLKTRKTSSEOAILPKLEICSDOGNLSPSRSTSVDDGLGSGYPLKDKSSGDOFLSV 75
DB 331 VPVKSRRKTTLEQPSVALKVEPVGEDLGSPPPADFAFEDYSSFOH-IRKGGFCDOYLAV 389

```
QY      76 PS-PFTWSKPKP 86
          |  | : | : |||
Db      390 PQHPYQWAKPKP 401
```

RESULT 15

ID AAE21563 standard; Protein; 716 AA.

AC AAE21563;

DT 16-JUL-2002 (first entry)

Human cytoplasmic nuclear factor of activated T cell (NF-ATc) protein #2.

KW Human; immunosuppressive; cytoplasmic nuclear factor of activated T cell;

KW NF-ATc; nuclear translocation.

05 Homo sapiens.

PN US6352830-B1.

PD 05-MAR-2002.

PF 15-JAN-1999; 99US-0232346.

PR 22-AUG-1991; 91US-0749385.

PR 20-SEP-1993; 93US-0124981.

PR 31-JUL-1995; 95US-05070332.

PA (STRD) UNIV LELAND STANFORD
PA
...

PI Cřabtree GR, Northrop JP,

DR WPI; 2002-314700/35.

XX
XX
T2004-4-10 2:33:00 PM

PT cytoplasmic NF-AT polypeptide

PT Translocation -

PS Claim 20; Column 81-84; 83pgs
 XY

The invention relates to a method comprising:

CC nuclear factor of activated

CC Identifying an immunosuppres

cc The present sequence is number
xx

sequence / 10 AA;

query malch	24.88;
Best local similarity	34.78;

malclies 23; conservative 1

09 1P LK R I S E D Q A A L P G K L E I C

DB 331 VPVKSRIITLEQPSVALKVEPV

PS-PEIWSKPKP 86

DD 390 P0HPYQWAKPKP 401

Search completed: April 11 2003

```
search completed: April 11, 2003,
Job time : 40 secs
```


Db 321 LGPAVFPPQYVEVDIDILKTRKRTISEDQAAILPGLKLELCSDDQGSLSPARETSTIDDELGSQ 380

QY 61 YPLKDDSGDQFLSVSPFTWMSKRP 86
||||| ||||||| ||||||| |||||||
Db 381 YPLKDDSGDQFLSVSPFTWMSKRP 406

```

RESULT 3
S45262
NF-AT component - human
C:Species: Homo sapiens (man)
C:Date: 10-Dec-1994 #sequence.revision 10-Nov-1995 #text.change 28-May-1999
C:Accession: S45262
R:Northrop, J.P.; Ho, S.N.; Chen, L.; Thomas, D.J.; Timmerman, L.A.; Nolan, G.P.; Admon,
Nature 369, 497-502, 1994
A:Title: NF-AT components define a family of transcription factors targeted in T-cell ac
A:Reference number: S45262; MUID:94261186; PMID:8202141
A:Accession: S45262
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-716 <NOR>
A:Cross-references: GB:U08015; NID:g500631; PIDN:AAA19601.1; PID:g500632

```

RESULT 4
JC5805
transcription factor NFATC - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 07-May-1999
C:Accession: JC5805
R:Pan, S.; Koyano-Nakagawa, N.; Tsuruta, L.; Amsahl, Y.; Yokota, T.; Morl, S.; Aral, N.
Biochem. Biophys. Res. Commun. 240, 314-323, 1997
A:Title: Molecular cloning and functional characterization of murine cDNA encoding trans
A:Reference number: JC5805; MUID:98049829; PMID:9388475
A:Accession: JC5805
A:Molecule type: mRNA
A:Residues: 1-718 <PAN>
A:Note: the sequences of residues 30-39 and 40-59 are interchanged in the authors' trans
C:Comment: This protein plays a role in immune and inflammatory response by regulating t
F:202-211,236-245,281-290/Region: Sp-box
F:684-687/Region: nuclear location signal

```
RESULT 5
G86253
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
;Accession: G86253
```

R.T. Theologis, E. Ecker, J.R. Palm, C.J. Federpiel, N.A. Kaul, S. White, O. Alon, Chiu, C.W., Chung, M.K., Conn, L., Conway, A.B., Conway, A.R., Cressy, T.H., Dewar,ansen, N.F., Hughes, B.B., Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziano, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallio,ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86253
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 15329 <STO>
A:Cross-references: GB:AE005172; MID:93157936; PIDN:AACT1619.1; GSPDB:GN00141
A:Genetics:
A:Map position: 1

RESULT 6
T34418
hypochemical protein F12F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34418
R:Fulton, B.; Woldmann, P.
Submitted to the EMBL data library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: 221521
A:Accession: T34418
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-3488 <FUL>
A:Cross-references: EMBL:U00023; PTDN:AAC25685.1; GSPPDB:GN00023; CESP:F12F3.3
A:Experimental source: strain Bristol N2, clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.3
A:Map position: 5
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3593/1

RESULT 7
 AB0821
 probable exported protein SMY2760 [imported] - *Salmonella enterica* subsp: *enterica* serovar Typhimurium
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhimurium
 A:Note: This species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
 #Accession: AB0821

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov A:Reference number: AB0502; PMID:11677608
 A:Accession: AB0821
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2130 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD0218.1; PID:g16503731; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY2760

Query Match 15.7%; Score 72; DB 2; Length 2130;
 Best Local Similarity 27.0%; Pred. No. 30;
 Matches 24; Conservative 13; Mismatches 38; Indels 14; Gaps 2;

Oy 5 PFPOYCVETDIPLTKRTKTSDDQAAILPGKLEICSDDOGNLSPSRETSV-----53
 Db 501 PLPYRYMELFDPAKNNK-GEHODAMNORVTVSSEDMRASSPHYTGVTVDNGAHLTL 559

Oy 54 --DDGLSGOYPLKNDSSGDOFLSYSPSPT 80
 Db 560 QHDSGCGETPRITVMPDDEGNELPFS 588

RESULT 8

JC7122
 protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)
 N:Alternate names: serine (threonine) protein kinase
 C:Species: Mus musculus (house mouse)
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001

R:Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.
 Biochem. Biophys. Res. Commun. 264, 449-456, 1999

A:Title: Activity and substrate specificity of the murine STK2 serine/threonine kinase

A:Reference number: JC7122; MUID:20001940; PMID:10529384

A:Accession: JC7122

A:Status: preliminary

A:Molecule type: mRNA
 A:Residues: 1-792 <HAY>
 A:Cross-references: GB:AJ223071; NID:g4138208; PID:g4138209

C:Genetics:
 A:Gene: MSTRK2L

C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol C:Keywords: phosphotransferase

Query Match 15.6%; Score 71.5; DB 2; Length 792;
 Best Local Similarity 44.7%; Pred. No. 11;

Matches 17; Conservative 2; Mismatches 14; Indels 5; Gaps 1;

Oy 39 SDDOGLNLSPKRSTSVDDGLGSGQ-----YPLKNDSSGDO 71
 Db 418 SDOPGNLLPRRSSDGGCGEGSELVPLPSNKKDQPPQ 455

RESULT 9

SS4395
 permease-like protein - Streptococcus pneumoniae (fragment)

C:Species: Streptococcus pneumoniae

C:Date: 27-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 15-Oct-1999

C:Accession: SS4395

R:Pearce, B.J.; Naughton, A.M.; Measure, H.R.
 Mol. Microbiol. 12, 881-892, 1994

A:Title: Peptide permeases modulate transformation in *Streptococcus pneumoniae*.

A:Reference number: SS4395; MUID:95020610; PMID:7523829

A:Accession: SS4395
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-643 <PEA>
 A:Cross-references: EMBL:L20556; NID:g516639; PIDN:AAA26952.1; PID:g516640

C:Genetics:
 A:Gene: p1pA

Query Match 15.3%; Score 70; DB 2; Length 643;
 Best Local Similarity 34.1%; Pred. No. 12;

Matches 28; Conservative 12; Mismatches 26; Indels 16; Gaps 6;

Oy 8 FOYCVETDIP-----LTKRK--TSEDQAAILPGKLEICSDDOGNLSPS--RETSDV-DGLG-58
 Db 19 FSYIYERDPOULNLTTFKAATANITSNVVDGLLE--NDRNGNVPMSAEDMSVSKDGLT 76

Oy 59 SOYPLKND-----SSGDOFLSV 75
 Db 77 YTYTIRKDAKWTSEGEYAAV 98

RESULT 10

B37953
 transcription regulator PAN-2 - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: B37953

R:German, M.S.; Blamir, M.A.; Nelson, C.; Moss, L.G.; Rutter, W.J.
 Mol. Endocrinol. 5, 292-299, 1991

A:Title: Two related helix-loop-helix proteins participate in separate cell-specific

A:Reference number: A37953; MUID:91246228; PMID:1710033

A:Accession: B37953

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA
 A:Residues: 1-649 <GER>
 C:Superfamily: human transcription factor 3

Query Match 15.3%; Score 70; DB 1; Length 649;
 Best Local Similarity 26.1%; Pred. No. 12;
 Matches 30; Conservative 6; Mismatches 17; Indels 62; Gaps 6;

Oy 27 QAAILPGKLEICSDDOGNLSPS-----RETSDVDDGLGSGQ-----60
 Db 123 QAGFLPGLGLSS--PPPLSPSGVSGSQYPSYSPNRRRAADSGLDTSKKYKYPG 180

Oy 61 -----YPLKNDSSGDOFL-----LSVPSPT-----TWSPK 84
 Db 181 LPSSVYP-----SSGDSYGRDAAYPSAKTPGSAVSPPYVADGSLHPSAELWSP 232

RESULT 11

G95177
 hypothetical protein SPI527 [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: G95177

R:Attell, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappel

son, T.; Hickey, E.R.; Holt, I.E.
 Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: G95177
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-652 <KUR>
 A:Cross-references: GB:AE005672; PIDN:AAK75616.1; PID:g14973017; GSPDB:GN00164; TIGR

A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SPI527

Query Match 15.3%; Score 70; DB 2; Length 652;
 Best Local Similarity 30.3%; Pred. No. 12;
 Matches 27; Conservative 13; Mismatches 31; Indels 18; Gaps 5;

Oy 2 SPAPFOYCVETDIPLTKRTKTSDDQAAILPGKLEICSDDOGNLSPSRE--T 51
 Db 19 FSYIYERDPOULNLTTFKAATANITSNVVDGLLE--NDRNGNVPMSAEDMSVSKDGLT 76

```

Db      29 STASTYNNVSSD-PSSLNTLAENKRAFTSDIVANLVGELLE-NDQXGNIIPLAEEDWT    85
Ox      52 SYVDDELGSQYPLKRD-----SSGDQFLSY    75
          |||   |   |   |   |   |   |
Db      86 VSQDLGYTYVKLRDKAKWFTSGEEYAPY    114

RESULT 12
E98044
hypothetical protein alb [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: E98044
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; F
e, R.; LeBlanc, D.J.; Lee, L.N.; Defkowitz, E.J.; Lu, J.; Matsushima, P.; Mahren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: E98044
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-652 <KUD>
C:Cross-references: GB:AEO07317; PIDN:AAL00186.1; PID:g15459032; GSPDB:GN00174
C:Genetics:
A:Gene: albB

```

	Query Match	15.3%;	Score 70;	DB 2;	Length 652;	
	Best Local Similarity	30.3%;	Pred. No. 12;			
	Matches	27;	Conservative 13;	Mismatches 31;	Indels 18;	Gaps 5;
Oy	2 SPAPPFOYCVETDIPKTRKTSEDOA-----ILPKLEICSDDDGNNLSPRE---T	51				
	: :					
Dd	-29 STASTTYNNVYSSD--PSSLNYLAENRAATSDIVANLVLDGLE--NDQYNIIPTSLAEPMW	85				
Oy	52 SVDDGLGSQYPLKRD-----SSGDQFLSY	75				
	: :					
Dd	86 VSQDLTYTYTKLRKDAAKFTSGEETYAPV	114				

```

RESULT 13
G97912
hypothetical protein alia [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: G97912
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Fea, R.; LeBlanc, D.-J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
Y. J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G97912
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-660 <KUR>
A:Cross-references: GB:AE07317; PIDN:AAK9131.1; PID:g15457683; GSPDB:GN00174
C:Genetics:
:Gene: alia

```

	Query Match	Similarity	Score	DB	Length
Best Local	28	15.3%	70	2	660
Matches	28	Conservative	12	Mismatches	26
				Indels	16
				Gaps	6

RESULT 14
 E95042
 hypothetical protein SP0366 [Imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: E95042
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Reed, T.D.; Peterson, S.; H
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
 l, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000, MUID:21357209, PMID:11463916
 A:Accession: E95042
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-661 <KUR>
 A:Cross-references: GB:AF005672; PIDN:AAK74534.1; PID:g14971835; GSPDB:GN00164; TIGR:
 A:Experimental source: Strain TIGR4
 C:Genetics:
 A:Gene: SP0366

```

Oy      8  FÖRVEFENDIP-----LKTR--TSEDDAALPKKLEICSDGDNLSPS--KETSYD-DGLG 56
        |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      37  FSTIYEYDPPNLLYLTAAKAAANTITSNVGDLE--NDRIGNEVPSMAEDMSVSKDGLT 94
        |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Oy      59  SÖYPLKAD-----SSGDÖFLSY 75
        |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db      95  YTYIRKDAWYTSSEGEYAAV 116
        |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

```

RESULT 15
B35816
transcription regulator Pan-1 - rat
N:Alternate names: Insulin enhancer-binding protein 1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 16-Jul-1999
C:Accession: B35816; A35487
R:Nelson, C.; Shen, L.P.; Meister, A.; Fodor, E.; Rutter, W.J.
Genes Dev. 4, 1035-1043, 1990
A:Title: Pan-1: a transcriptional regulator that binds chymotrypsin, insulin, and Ap-4
A:Reference number: A35816; MUID:90346284; PMID:2200736
A:Accession: B35816
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-638 <NEU>
A:Cross-references: GB:X62323; NID:G35277; PIDN:CMAA4199.1; PID:q1045195
R:Shibasaki, Y.; Sakura, H.; Takaku, F.; Kasuga, M.
Biochem. Biophys. Res. Commun. 170, 314-321, 1990
A:Title: Insulin enhancer binding protein has helix-loop-helix structure.
A:Reference number: A35487; MUID:90321246; PMID:2196879
A:Accession: A35487
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 297-377,'O',378-638 <SHI>
C:Superfamily: human transcription factor 3
C:Keywords: alternative splicing; DNA binding; nucleus; transcription regulation

Query Match	15.1%	Score 69.5;	DB 2;	Length 638;
Best Local Similarity	25.7%	Pred No. 14;		
Matches	29;	Conservative	5;	Mismatches 22; Indels 57; Gaps 5;
Oy	27	QAALPGKLEICSDDGNSLPS	-----	RETSYDDIGSGOYPLK-----64
Db	114	QAAGLPGLPELGLIS--	PGPLSPSGYKSSSQYYTSPSPNRRRADGGGLADLPQPKVRKVP	171
Oy	65	-----KDSGGDGF	-----	LSVSPF-----TWSKP84

Mon Apr 21 09:44:41 2003

us-09-550-115a-2.rpr

Page 5

Db 172 GLPSSVPPSSGDNYSRDATAYPSAKTPSSAYPSPFYVADCSLHPSAELMSPP 224

Search completed: April 11, 2003, 14:23:42
Job time : 21 secs

WEST Search History

DATE: Monday, April 21, 2003

<u>Set Name</u> side by side	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u> result set
<i>DB=USPT,PGPB,JPAB,EPAB; PLUR=YES; OP=OR</i>			
L17	L16 and (l8 or l9 or l10)	19	L17
L16	L15 and l2	258	L16
L15	L14 and l1	1224	L15
L14	fk506 or fk-506 or "fk 506"	2736	L14
L13	fk506 or fk-506 or :fk 506	101057	L13
L12	L11 and (l8 or l9 or l10)	24	L12
L11	l1 and l2	368	L11
L10	((435/21)!.CCLS.)	625	L10
L9	((435/4)!.CCLS.)	4330	L9
L8	((530/324)!.CCLS.)	3223	L8
L7	"calcineurin binding"	54	L7
L6	calcineurin.ti.	19	L6
L5	(l3 or l4) and l1	3	L5
L4	"ARAI; KEN-ICHI"	28	L4
L3	"LIU; JIE"	40	L3
L2	nfat\$ or "nf at" or nf-at\$	644	L2
L1	calcineurin or cn	88767	L1

END OF SEARCH HISTORY

